



The University of
Nottingham

UNITED KINGDOM • CHINA • MALAYSIA

**Investigation of protein arginine methyltransferase
bisubstrate inhibitors**

Alaa AL-Noori

Supervisors:

Dr James Dowden

Dr Ingrid Dreveny

School of Chemistry

**Thesis submitted to the University of Nottingham for the degree
of Doctor of Philosophy**

September 2019

Declaration

I hereby declare that this thesis is my own work and that it has not been submitted anywhere for any other degree or award. The work presented herein is my own and where other sources of information have been used, they have been duly acknowledged.

Alaa Alnoori

Dr. James Dowden

Dr. Ingrid Dreveny

**This thesis is dedicated to the spirit of my father and my brother
Riadh Alnoori in appreciation of their many contributions to my
well-being before they passed away.**

Abstract

Protein arginine methylation is a significant post-translational modification which is catalysed by protein arginine methyltransferases (PRMTs). PRMTs catalyse the transfer of methyl group from the cofactor *S*-adenosyl-L-methionine (SAM) to arginine residues which are present in proteins. Dysregulation of these enzymes has been associated with cellular diseases such as prostate cancers, cardiovascular disease, and spinal muscular atrophy.¹

This thesis concerns the design and synthesis of SAM analogues, primarily as PRMT inhibitors, and evaluation of their activities against a selection of these enzymes. Bisubstrate inhibitors were designed to include SAM structure as well as those that mimic the peptidyl arginine substrate. The design of the inhibitors was based on potent and selective inhibitors of PRMTs, which had already been developed within the research group.^{2,3} The sulfonium centre of SAM was replaced with the more stable nitrogen atom to help produce robust SAM analogues which possess better chemical stability. In this thesis, we introduce a 2-aminopyridine as an isosteric replacement for arginine. Temporary protection of the pyridine nitrogen by complexation with borane was found to be essential for high target inhibitor yields. Additionally, a route for the development of SAM co-factor analogues featuring hydrophobic esters at the 2',3'-hydroxyls of adenosine was also developed which may enable these compounds to pass through the cell membrane.

PRMT1 and CARM1 were produced which carried a His-tag at the N-terminus, and this His-tag was used to purify the proteins. Preliminary biological evaluation of

the inhibitors showed that AzaSAM analogues, which included a 2-amino pyridine as a guanidine isostere were found to be potent inhibitors and were somewhat selective towards CARM1 over PRMT1.

Human CARM1 domain (residues 135-479) was cloned and produced as a maltose-binding protein (MBP) fusion protein with a C-terminus His-tag. Suitable conditions were determined for cleavage of the MBP tag and purification of CARM1 from the resulting mixture.

In order to improve the yield of the CARM1 domain (residues 135-479), it was cloned into a pRSF-13 vector as a 2GKG fusion protein which possessed a TEV cleavage site to allow subsequent removal of the 2GKG tag. CARM1 was successfully co-crystallised with two inhibitor compounds described in this thesis.

Human PRMT6 (residues 28-375 and 39-375-6×His) was cloned and produced with a soluble tag at the N-terminus and a His-tag at the C-terminus. PRMT6 (residue 28-375-6×His) was separated from contaminating chaperonin proteins and was shown to be activated *via* an *in vitro* methylation assay.

Moreover, full-length human PRMT2 (residues 1-431-6×His) and a slightly shorter PRMT2 construct (residues 95-431-6×His) were produced and purified using the His-tag present at the C-terminus.

Acknowledgements

I would like to express grateful respect to my worthy supervisor's Dr James Dowden and Dr Ingrid Dreveny for giving me this great opportunity join their research group and offering continuous guidance and encouragement throughout my PhD study. I am also very grateful for their insightful comments on this thesis and kind advice under their supervision. I would like to thank Dr Clare Davies for supplying the template plasmids of PRMTs.

I am grateful to Dr Emma Gunnell for carrying out *in vitro* methylation assays experiments on my behalf using these materials as indicated throughout the thesis. I would like to express my sincere thanks to Dr Tim Birkinshaw, Usama, Rasha and Simon for their support and help whenever I needed. I am thankful to my colleagues in both groups, past and present for providing a comfortable and co-operative working environment. I should like to thank the technical staff in the School of Chemistry and School of Pharmacy, NMR staff Kevin, Shaz, Graham and mass spectrometry analysis Mick and Ben for technical assistance.

I would like to say a huge thank you to my family, special for my dad and to my mum, my uncle Qasim, and Dr Saad Al-Hashimi and for all their support, love and advice during this time. I couldn't have done it without you all. Finally, I express my warmest gratitude and love to my wife, Entedhar Rabiaa and her father Dr Kadhum and her mother Sabeeha for their understanding and support during the research period. I would also like to express my profound gratitude to my little boys Riadh, Rayan and Yazn for making me so happy with their cute smile and engaging me a lot at home.

Table of contents

Abstract.....	iii
Acknowledgements.....	v
Table of contents.....	vi
Abbreviation and symbols.....	xiii
Crystallographic terms	xvii
1. Chapter 1 Introduction	1
1.1. <i>S</i> -Adenosyl methionine and methyltransferases.....	1
1.1.1. Synthesis and metabolism of SAM	2
1.1.2. Chemical stability of SAM	5
1.2. Protein arginine methylation	7
1.2.1. Arginine residue	8
1.2.2. The proposed catalytic mechanism	11
1.2.3. PRMT substrates and product specificity	14
1.3. Structure and functional analysis of PRMTs.....	15
1.4. Protein arginine methylation and diseases	17
1.5. PRMT1 and CARM1	18
1.5.1 Biological roles of PRMT2 and PRMT6	19
1.5.2. PRMT6 structure	20
15.3. Homodimerisation of PRMT6	22
1.5.4. Signature motifs of PRMT6	23
1.5.5. Co-factor binding mode of PRMT6.....	24
1.5.6. Conformations of SAM when in the active site of methyltransferases	28

1.6. Inhibitors of PRMTs.....	29
1.6.1. Bisubstrate inhibitors.....	32
1.7. Project aims	38
1.8. Prototype PRMT inhibitors.....	39
1.9. Preparation of hydrophobic version for SAM analogue.....	41
2. Chapter 2 Chemistry results and discussion	42
2.1. Aims and objectives.....	42
2.1.1. Inhibitor synthesis	43
2.1.2. SAM analogues containing hydrophobic isostere 2-aminopyridine....	48
2.1.3. Nitrogen SAM analogue containing guanidine.....	54
2.2. SAM analogue containing butyl ester on 2' and 3' of adenosine	56
2.2.1. Synthesis of the model compound	57
3. Chapter 3 Optimising the production, purification and crystallisation of human PRMT1 and CARM1	66
3.1. Choice of constructs	66
3.2. PRMT1: purification and crystallisation.....	68
3.2.1. Choice of construct	68
3.2.2. PRMT1 production and purification.....	69
3.2.3. Evaluation of selected compounds as inhibitors of PRMT1 by <i>in vitro</i> methylation assay.....	72
3.2.4. Isothermal titration calorimetry (ITC) with inhibitor compounds.....	75
3.2.5. Crystallisation assays with hPRMT1 and bisubstrate inhibitors	77
3.3. CARM1: purification and crystallisation studies.....	78

3.3.1. Construct designed of CARM1	78
3.3.2. Purification of MBP-CARM1 (residues 135-479-6xHis).....	78
3.3.3. Improving the yield of CARM1 production using an alternative tag ..	81
3.3.4. CARM1 with a cleavable 2GKG soluble tag.....	82
3.3.5. <i>In vitro</i> methylation assay to test the activity of compounds with CARM1	84
3.3.6. Isothermal titration calorimetry (ITC) and IC ₅₀ for CARM1.....	87
3.3.6.1. Crystallisation of CARM1 with small molecule inhibitors	91
3.3.6.2. Data collection and structure determination.....	93
3.3.6.3. Bisubstrate inhibitors binding to the active site of CARM1	96
3.4. CARM1 with a non-cleavable 2GKG soluble tag.....	100
3.4.1. ITC analysis for CARM1 with non-cleavable 2GKG soluble tag	102
3.4.2. Crystallisation	104
4. Chapter 4 PRMT6.....	106
4.1. Rationale and construct design for PRMT6 (EC: 2.1.1.319).....	106
4.2. Purification of human PRMT6.....	108
4.2.1. PRMT6 construct design.....	108
4.2.1.1. Production and purification of human PRMT6	109
4.3. Utilising an N-terminus MBP tag optimised for PRMT6 to improve protein purification	111
4.3.1. PRMT6 catalytic domain-soluble tag fusions.....	114
4.3.2. Test production of PRMT6 with fusion tag 2GKG and1BKR.....	115
4.3.3. Utilising an N-terminus 1BKR tag optimised for hPRMT6	116

4.4. 1BKR-PRMT6-His expression in Arctic Express cells	120
4.4.1. Western blot used for 1BKR-PRMT6-His	122
4.4.2. Mass spectrometry	123
4.4.3. Removal of co-purifying chaperonin Cpn 60/10 ¹⁰⁹	124
4.4.4. Autoinduction at a higher cell density increases protein yield ¹⁰⁶	126
4.4.5. Isothermal titration calorimetry (ITC)	128
4.4.6. Evaluation of human 1BKR-PRMT6 catalytic activity	130
4.4.7. Radiometric (³ H) activity assays	131
4.4.8. Linear range determination for 1BKR-PRMT6-His	133
4.5. Test inhibition of small molecule inhibitors against hPRMT6	133
4.6. Crystallisation	136
4.6.1. Data collection evolution of diffraction	138
4.7. Summary	141
5. Chapter 5 Characterisation of human PRMT2	142
5.1. Full-length and catalytic domain constructs of human PRMT2	142
5.1.1. Attempted production and purification of FL-PRMT2 construct	145
5.1.2. FL-PRMT2 production and purification in (TF) fusion protein	146
5.2. Attempted production and purification of the catalytic core domain of PRMT2	150
5.2.1. PRMT2 construct design	150
5.2.2 Purification of PRMT2 (residues 95-431-6×His) with soluble tags	150
5.2.3. PRMT2 in fusion with 1BKR soluble tag	151

5.2.4. Analysis of 1BKR-PRMT2 by cleavage time-courses.....	154
5.2.5. Western blot	155
5.3. Purification of repressor splicing factor 1 (RSF1).....	156
5.3.1. <i>In vitro</i> methylation assay.....	161
5.3.2. Co-production of 1BKR-PRMT2-His with RSF1 or H3.....	162
5.4. Crystallisation of 1BKR-PRMT2-His	167
5.4.1. Data collection for PRMT2.....	168
5.4.2. SDS-PAGE analysis western blot for a crystal of 1BKR-PRMT2-His ..	169
5.5. Summary.....	171
6. Chapter 6 Conclusion and future work.....	172
7. Chapter 7 chemistry experimental	177
7.1. Experimental.....	179
7.2. Materials and methods of biological chapters	225
7.2.1. Chemical components	225
7.2.2. Buffers pH adjustment.....	225
7.2.3. Biological material	227
7.3. Molecular biology techniques.....	227
7.3.1. Sterilisation and sterile technique	227
7.3.2. Agar plates	228
7.3.3. Plasmid DNA purification	228
7.3.4. DNA agarose gel electrophoresis	229
7.3.5. Glycerol stocks	229
7.3.6. Generation of construct.....	230

7.4. Polymerase chain reaction.....	233
7.4.1. Restriction enzyme digestion	235
7.4.2. Phosphatase treatment	235
7.4.3. Ligation.....	236
7.4.4. Transformation.....	237
7.4.5. Cracking cells.....	237
7.4.6. DNA sequencing	238
7.4.7. Site-directed mutagenesis.....	239
7.5. BL21-CodonPlus (DE3)-RIL competent cells (Agilent), LOBSTR BL21 (DE3)- RIL and, Arctic Express competent cells.....	240
7.6. Gene expression and protein production.....	241
7.7. Protein purification	244
7.7.1. General protein purification by FPLC.....	244
7.7.2. Cell lysis.....	244
7.7.3. Nickel affinity chromatography	245
7.7.4. Gel filtration.....	246
7.8. High-density shaking cultures by auto-induction media.....	248
7.8.1. Media preparation	248
7.8.2. Gene expression and cell lysis.....	249
7.8.3. A single step separation of co-purified chaperonins.....	250
7.9. Protein analytical techniques	251
7.9.1. SDS-PAGE	251

7.9.2. Gel preparation	251
7.9.3. Sample preparation.....	252
7.9.4. Gel staining	253
7.10. Western blot detecting His-tagged proteins.....	253
7.11. Crystallisation	254
7.11.1. Sample preparation and crystallisation	255
7.11.2. 96-well screens.....	256
7.11.3. Optimisation screens using 24-well plates	256
7.11.4. Crystal fishing and freezing.....	257
7.12. Biophysical techniques	257
7.12.1. UV spectroscopy	257
7.12.2. Isothermal titration calorimetry (ITC)	258
8. Chapter 8 References	260
9. Appendix.....	270
9.1. <i>In Vitro</i> radiometric methylation assay.....	270
9.2. X-Ray diffraction, data collection, structure determination and refinement	271

Abbreviation and symbols

2GKG	Receiver domain of <i>Myxococcus xanthus</i> social motility protein FrzS (PDB code 2GKG)
1BKR	The calponin homology domain from human β -spectrin (referred to as the PDB code: 1BKR)
aa	Amino acid
aDMA	Asymmetric dimethylarginine
Ado	Adenosine
APS	Ammonium persulfate
ATP	Adenosine-5'-triphosphate
Ac	Acetate
ACOH	Acetic acid
BL21cp	BL21-CodonPlus (DE3)-RIL competent cells (Agilent)
Boc	<i>tert</i> -Butoxycarbonyl
cDNA	Complementary DNA
cspA	Cold shock protein A
CARM1	Co-activator-associated arginine methyltransferase 1
COSY	Homonuclear correlation spectroscopy
DIAD	Diisopropylazodicarbonate
Da	Dalton
DPM	Disintegrations per minute
DEPTQ	Distortionless enhancement by polarisation transfer with quaternaries
DEAD	Diethyl azodicarboxylate
DLS	Diamond light source
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DMP	Dess-Martin periodinane
DNA	Deoxyribonucleic acid
DP	Differential power
DIPEA	<i>N,N'</i> -Diisopropylethylamine
DTT	Dithiothreitol

dNTP	Deoxynucleotide acid
dH₂O	Double distilled water
dNTPs	Deoxynucleotide triphosphate
ESC	Embryonic stem cells
EtOH	Ethanol
ESI	Electrospray ionisation
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
Et₃N	Triethylamine
ExPASy	Expert protein analysis system
<i>E. coli</i>	<i>Escherichia coli</i>
FL	Full length
FL-PRMT2	Full length protein arginine methyltransferase 2
FL-PRMT6	Full length protein arginine methyltransferase 6
FPLC	Fast protein liquid chromatography
GAR	Glycine-arginine-rich
GST	Glutathione S-transferase
GB	B1 domain of streptococcal protein G (PDB code 3GB1)
h	Hour(s)
His	Histidine
HF	High Fidelity
HRMS	High-resolution mass spectrometry
hnRNP	Heterogeneous nuclear ribonucleoproteins
HSQC	Heteronuclear single-quantum correlation
IC₅₀	Half maximum inhibitory concentration
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
Kb	Kilobase
KDa	KiloDalton
Kd	Dissociation constant
LB	Lysogeny broth
L	Litre(s)
MBP	Maltose binding protein

MMA	Monomethylarginine
mCARM1	Coactivator-associated arginine methyltransferase (mouse)
mRNA	Messenger ribonucleic acid
Me	Methyl
MW	Molecular weight
mAU	Milli absorbance units
Min	Minutes
MWCO	Molecular weight cut off
MRC	Medical research council
NMR	Nuclear magnetic resonance
NB	Nova blue competent cells
NEB	New England Biosciences
OD₆₀₀	Optical density at 600 nM
PACT	pH, Anion, Cation crystallisation suite
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PKMTs	Protein lysine methyltransferases
PRMTs	Protein arginine methyltransferases
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PGM	Proline/glycine/methionine/arginine-rich motif
PH	Plextrin homology
ppm	Parts per million
PPh₃	Triphenylphosphine
RNA	Ribonucleic acid
rpm	Revolutions per minute
rcf	Relative centrifugal force
r.t.	Room temperature
RSF1	Repressor splicing factor 1
RONN	Regional order neural network
SDS PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
sDMA	Symmetric dimethylarginine

SAH	S-Adenosyl homocysteine
SAM	S-Adenosyl-L-methionine
SET	Su(var)3-9, enhancer-of-zeste and trithorax
SH3	SRC homology 3 domain
TBST	Tris-buffered saline with TWEEN-20
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid (TFA) neat liquid
T_m	Melting temperature
TBAI	Tetrabutylammonium iodide
TBME	<i>tert</i> -Butyl methyl ether
TEV	Tobacco etch virus
TF	Trigger factor
THF	Tetrahydrofuran
TsOH	p-Toluenesulfonic acid
tRNA	Transfer RNA
Tris	Tris (hydroxymethyl) aminomethane
TEMED	<i>N,N,N',N'</i> -Tetramethylethylenediamine
TIM	Triosephosphateisomerase
I.M.A.G.E.	Integrated molecular analysis of genomes and their Expression Consortium
JCSG+	Crystallisation screen designed by Joint Centre for Structural Genomics
PHENIX	Python-based hierarchical environment for integrated crystallography
V/v	Volume-to-volume
W/v	Weight-to-volume
U	Units
UV	Ultraviolet
ZnF	Zinc finger
zPRMT2	Protein arginine methyltransferase 2 (zebrafish)
ΔG	Gibbs free energy
ΔH	Enthalpy
ΔS	Entropy

Crystallographic terms

Å	Angstrom
a, b, c	Real space unit cell edge vectors
a*, b*, c*	Reciprocal space unit cell dimensions
α, β, γ	Real space unit cell angles
AU	Asymmetric unit
CC 1/2	Pearson correlation coefficient between random half-datasets
h,k,l	Miller indices/Reciprocal lattice points
I	Intensity
λ	Wavelength
MR	Molecular replacement
NCS	Non-crystallographic symmetry
R	Crystallographic R-factor
R_{free}	Free R-factor
R_{merge}	Linear merging R factor
R_{pim}	Precision-indicating merging R factor
TLS	Translation/Libration/Screw
x,y,z	Real space coordinates

1. Chapter 1 Introduction

1.1. S-Adenosyl methionine and methyltransferases

SAM-dependent methyltransferases play an important role in the regulation of gene transcription in both prokaryotic and eukaryotic cellular processes. In the majority of human cells excluding gametes, the genome is located within 23 pairs of chromosomes, where the constituent DNA is wound around histone proteins to form a complex known as chromatin. Two copies of the four main histones (H2A, H2B, H3, and H4) combine to form repeating units called nucleosomes, each forming a fibre of the chromatin DNA-protein complex.⁴ Histones are small basic proteins which possess a flexible N-terminus, this region is rich in arginine and lysine residues and are termed "histone tails". These residues are positively charged at physiological pH and can interact with negatively charged phosphate groups of DNA. Various post-translational modifications of these residues can occur, such as methylation, acetylation, phosphorylation, sumoylation, ubiquitination, and glycosylation which modulate these interactions allowing regulation of chromosome structure and thus DNA regulation or repair.⁴

Biomethylation plays a key role in the regulation of a number of cellular processes in both prokaryotic and eukaryotic cells. Arginine methylation is of particular interest; this form of methylation is a common post-translational modification in histone and non-histone proteins and has been estimated to affect 0.5 % of all arginine residues.⁵ In this process, methyl groups are transferred from a cofactor, S-Adenosyl-L-methionine (SAM), to arginine residues within substrates, such as histone proteins. This process is catalysed by arginine methyltransferases (PRMTs,

EC 2.1.1.125), which were first identified over 45 years ago.⁶ These enzymes methylate many target proteins such as histones H3 and H4, and other proteins.⁷

1.1.1 Synthesis and metabolism of SAM

S-Adenosyl-L-methionine was first reported by Cantoni in 1952 as a biological sulfonium compound.⁸ SAM is considered the second most important co-factor after adenosine-5'-triphosphate (ATP) in the human body.⁹ SAM acts as a source of methyl groups for methyltransferases, but SAM also offers versatile reactivity as discussed later.¹⁰

SAM biosynthesis from methionine and adenosine 5'-triphosphate (ATP) is catalysed by methionine adenosyltransferase, while the enzyme tripolyphosphatase performs simultaneous hydrolysis of the tripolyphosphate by-product (figure 1-1).¹¹

PRMTs catalyse methylation *via* a presumed S_N2-like mechanism involving the direct transfer of a methyl group from SAM to a nucleophilic atom of the substrate leading to a co-product S-Adenosyl-L-homocysteine (SAH). This co-product is a competitive inhibitor of methyltransferases.^{10,12} S-Adenosyl-L-homocysteine hydrolase prevents the accumulation of SAH by hydrolysing to adenosine and homocysteine. Homocysteine can then either be converted into glutathione or transported out of the cell.¹⁰ Homocysteine can also be methylated to generate methionine by the action of methyltetrahydrofolate-homocysteine methyltransferase (Figure 1-1).¹⁰

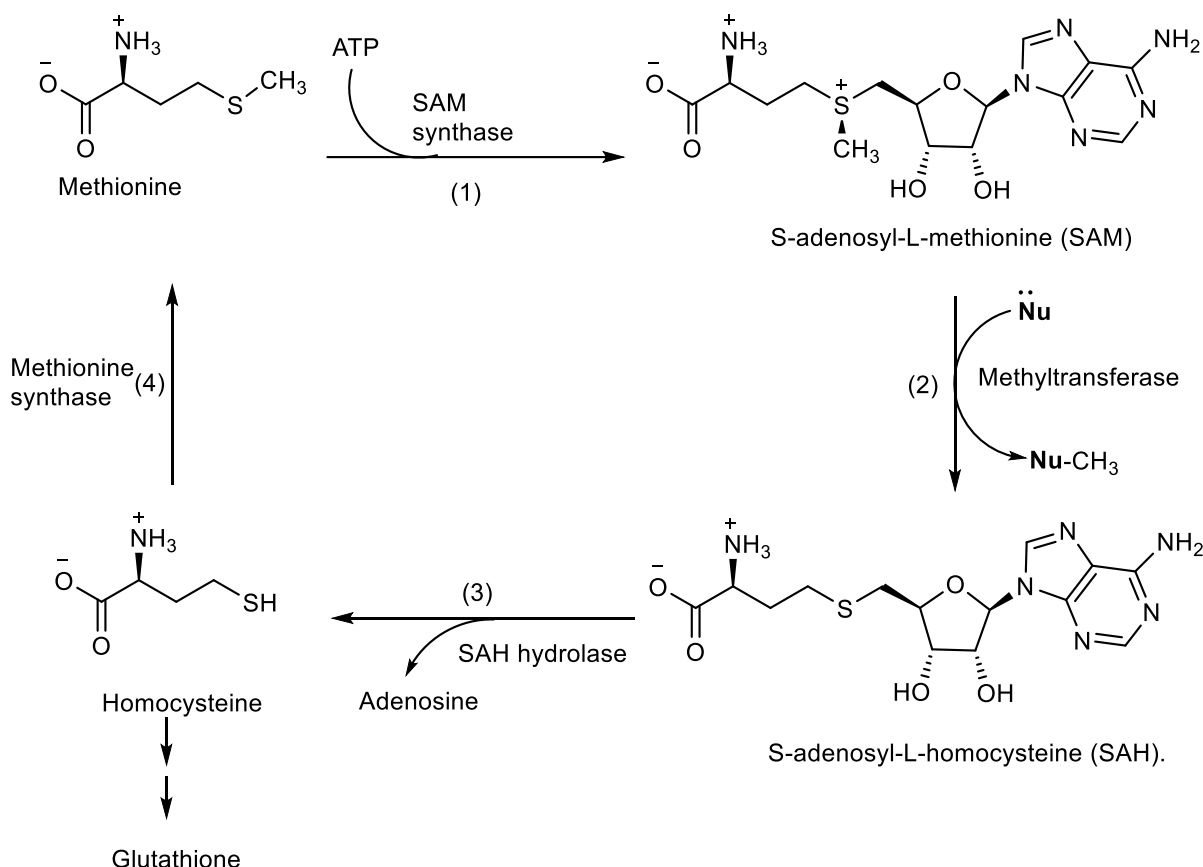


Figure 1-1: Biosynthesis of SAM within the cell. Methyl transfer of the methyl group from SAM is catalysed by enzyme methyltransferases and SAH's metabolism.¹⁰

SAM has many versatile roles; these roles generally arise from the electrophilic character of carbon centres which are adjacent to the positively charged sulfur (Figure 1-2).¹³ Generally, post-translational methylations may occur at any nucleophilic nitrogen (N-methylation), oxygen (O-methylation), or to lesser extent carbon (C-methylation) or sulfur (S-methylation) which are present within DNA, RNA and proteins or metabolites.¹⁰ For example, cyclopropane-fatty-acyl-phospholipid synthase affects methylation of the double bond of unsaturated fatty acids, leading to cyclopropanation (Figure 1-2).¹⁰

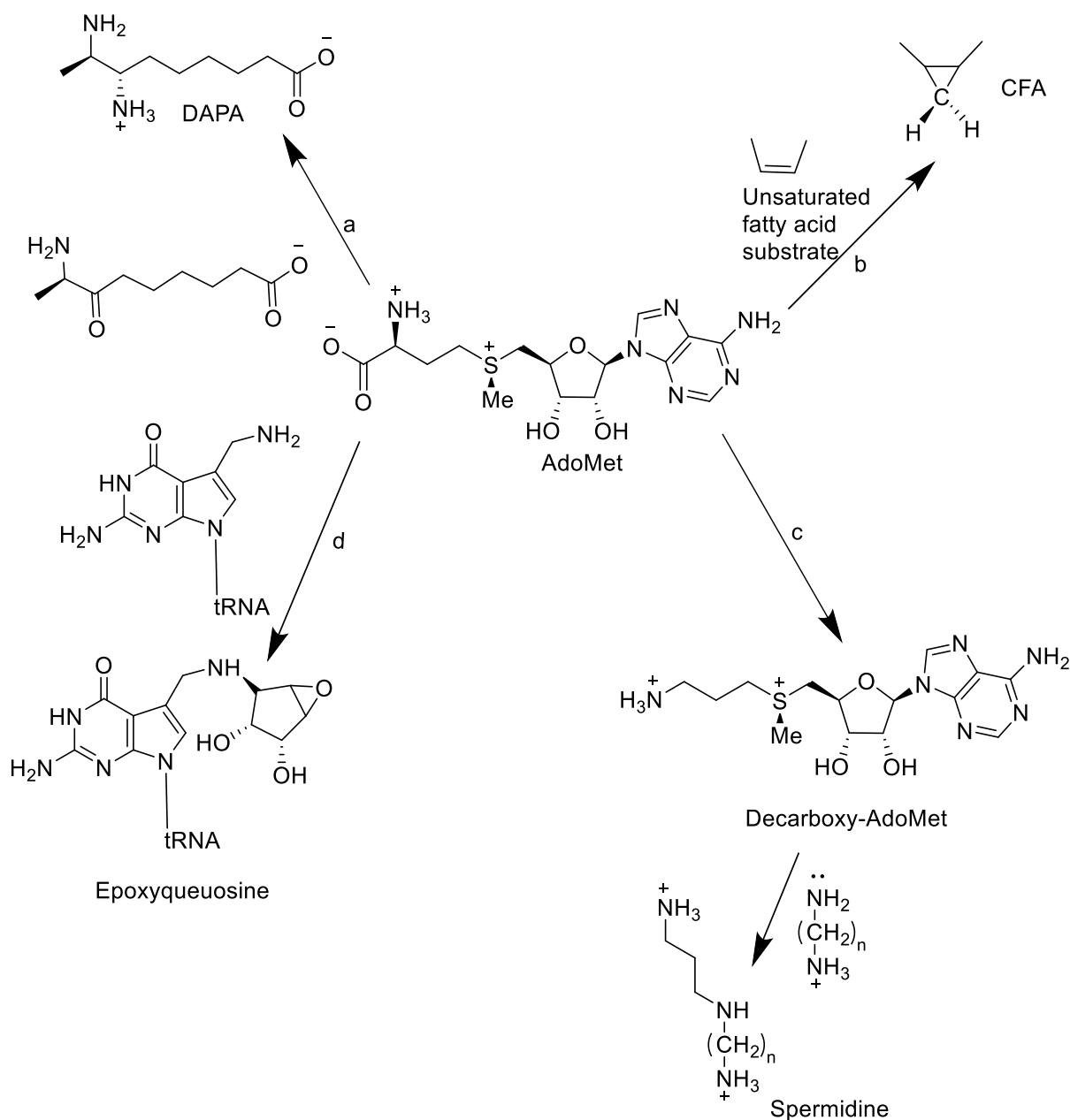


Figure 1-2: Example of AdoMet as a source of chemical groups functional.

A) amines (synthesis of 7, 8 diaminoperlagonic acids), B) methylene group (synthesis of cyclopropyl fatty acids), C) aminopropyl groups (synthesis of ethylene and polyamines), D) ribosyl groups (synthesis of epoxyqueuosine).

1.1.2. Chemical stability of SAM

The chemistry of *S*-Adenosyl-L-methionine (SAM) is dominated by the positively charged sulfonium centre which can possess either ionic or radical reactivity.¹³ SAM is highly unstable, and its decomposition which is pH-dependent;¹⁴ for example, the carboxylate anion can undergo internal attack by the methionine-derived γ -carbon, leading to 5'-methylthioadenosine (MTA) **1** and homoserine lactone (HSL) **2** (Figure 1-3).^{13,12}

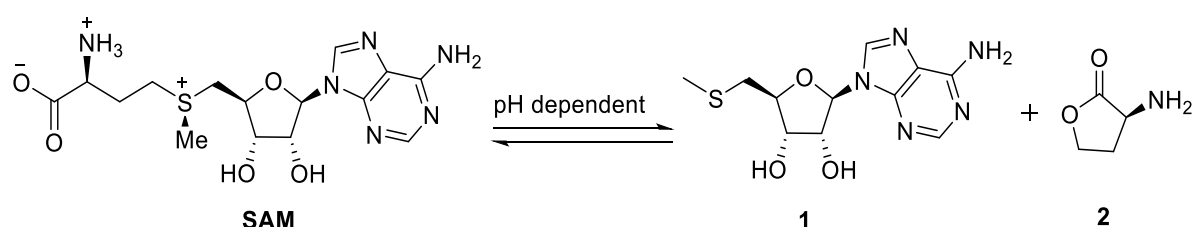


Figure 1-3: SAM hydrolysis into 5'-methylthioadenosine (MTA) and homoserine lactone (HSL).¹³

Alternatively, depurination of SAM may occur in an aqueous acidic solution, leading to *S*- ribosylmethionine **3** and adenine **4** (Figure 1-4).¹³

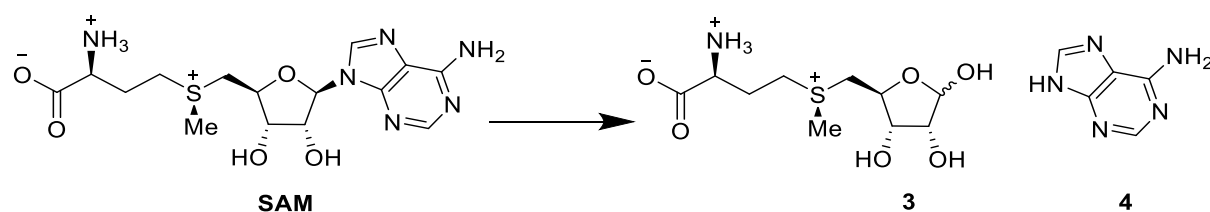


Figure 1-4: SAM also can hydrolysis into *S*-ribosylmethionine and adenine.

The sulfonium centre is itself chiral, and this has been reported to be essential for biological activity as a co-substrate of SAM.¹⁵ However, the sulfonium centre undergoes slow pyramidal inversion (Figure 1-5), such that commercial preparations of this molecule produces two epimers.¹⁴ The (*R*, *S*)-SAM is a potent inhibitor of methylases and constitutes 3 % of the total SAM found in mouse liver, while the (*S*, *S*)-SAM epimer is the biologically active form.¹⁴ It has been demonstrated that the configurational stability of SAM could be enhanced by maintaining acidic conditions (pH 2-3.5), low temperature (-20 or -70 °C) and using large non-nucleophilic counterions.^{14,13}

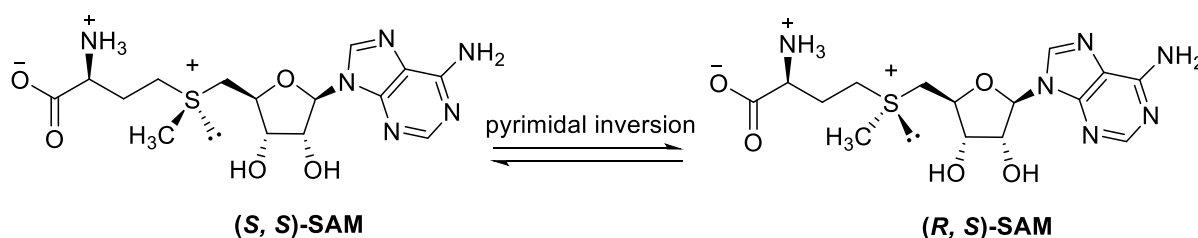


Figure 1-5: Inversion of the chirality at the sulfonium centre. SAM can exist as epimers; (*S*, *S*)-SAM active and (*R*, *S*)-SAM not active.

A viable strategy for the study of enzymes that can utilise SAM is to develop inhibitors which mimic this co-substrate. However, it is highly desirable to replace the sulfonium centre with a more stable functional group. Nitrogen is often used to replace sulfur since this element is still able to support a positive charge upon protonation (but generally undergoes rapid pyramidal inversion). The resulting tertiary amine centre is less reactive and provides for more robust SAM analogues with better chemical stability under physiological conditions.¹³ A further advantage of using nitrogen is that it is able to support functionality that can be designed to occupy the binding site of the other co-substrate (protein or peptide), resulting in bisubstrate inhibitor designs. This strategy has been used to develop inhibitors of

several human methyltransferases, such as histone methyltransferases SET7/9, indole *N*-methyltransferase, and DNA methyltransferases.¹⁶ The Dowden group has synthesised and evaluated bisubstrate inhibitors of PRMTs that incorporates functionality designed to resemble both the arginine containing co-substrate of PRMTs substrates and that of SAH and will be described in more detail later in this chapter.^{2,3}

1.2. Protein arginine methylation

Most studies have focused on the nine members of the PRMT family which have been identified in mammalian cells.¹⁷ There are three dominant methylated arginine products, which include monomethyl arginine (MMA), asymmetric dimethylarginines (aDMA), and symmetric dimethylarginines (sDMA), the last two terms are dependent up on which nitrogen of the guanidinium group carries the second methyl group (Figure 1-6).¹⁷ PRMTs are classified into four types, denoted as I to IV depending on the pattern of guanidine alkylation observed on their products. ^{18,19} Type I PRMTs (PRMTs 1, 2, 3, 4, 6, and 8 are examples) catalyse the formation of asymmetric dimethylarginine; while type II PRMTs (PRMTs 5, and 7) produce symmetric dimethylarginine, type III PRMTs catalyse the production of MMA observed for specific substrates (*e.g.*, PRMT 7), while the exact enzymatic activity of PRMT 9 methylation pattern has not been determined yet and type IV PRMTs produces δ -*N*-monomethylarginine as this enzyme catalyses monomethylation at the internal (or δ) guanidine nitrogen atom (Figure 1-6).^{19,17}

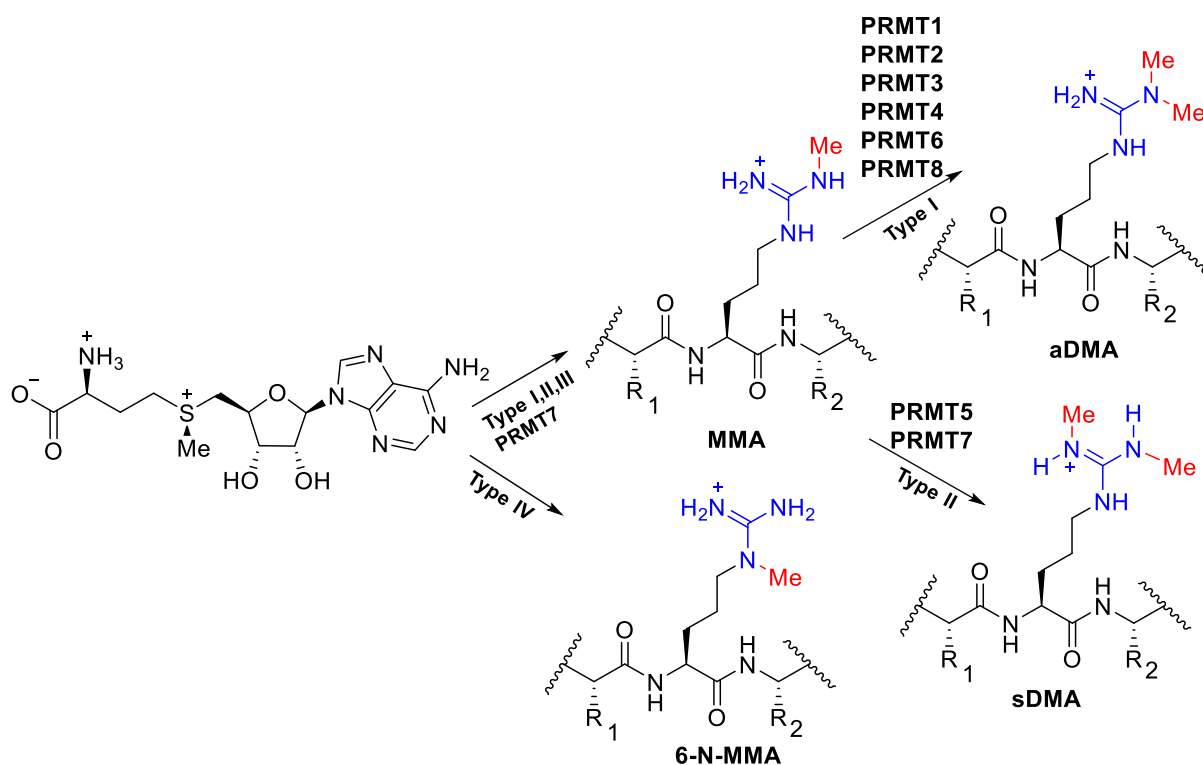


Figure 1-6: Arginine methylation patterns deposited by PRMTs.

The PRMTs catalysing methylation of arginine are categorised into four types including Type I which produces MMA and aDMA, and Type II which generates MMA and sDMA, Type III only generates MMA and Type IV generates 6-N-MMA, respectively.

1.2.1. Arginine residue

The guanidine group of arginine residues contains three nitrogens which can potentially be methylated (Figure 1-7). Methylation of arginine residues can be divided into three types according to the methylation pattern present on the target arginine in mammalian enzymes.²⁰ The first type of methylation is ω -N-monomethylarginine (MMA) with the resulting conformations divided into *anti*-MMA and *syn*-MMA. Subsequent methylation gives two types of isomers, the first being ω -N,N-asymmetric dimethylarginine (aDMA) and the second is ω -N,N'-symmetric dimethylarginine (sDMA) which is divided into *anti-anti*-SDMA and *anti-syn*-SDMA conformers (Figure 1-7).²⁰

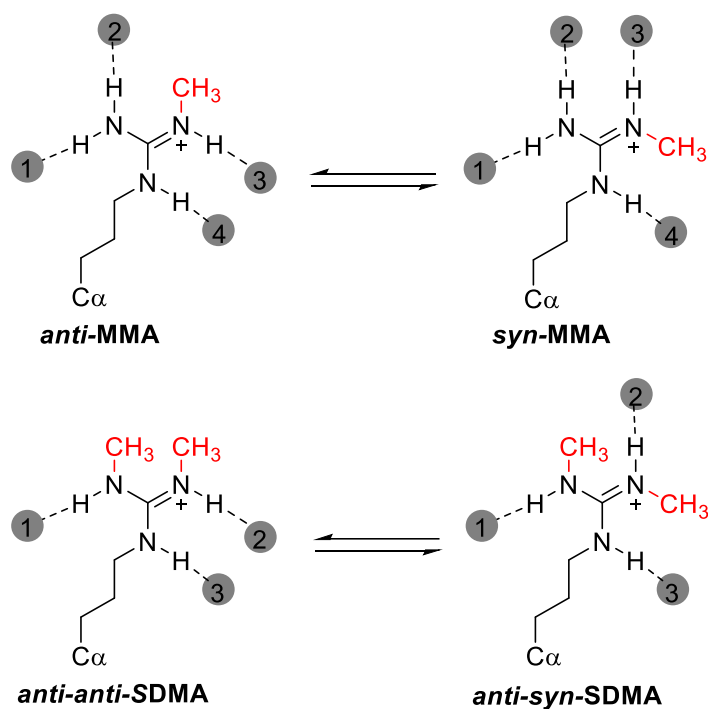


Figure 1-7: conformational exchange of methylated arginine's
Methyl groups on the arginine residue are marked in red, whereas hydrogen-bonding donor sites are highlighted in grey.

Arginine bears three guanidine nitrogen atoms which can be labelled based on their position, the two nitrogen's present at the terminus are designated as "u" and the internal one designated as "d". The resonance between the three nitrogen's gives a bond order of 1.33 (Figure 1-8).^{20,21,22}

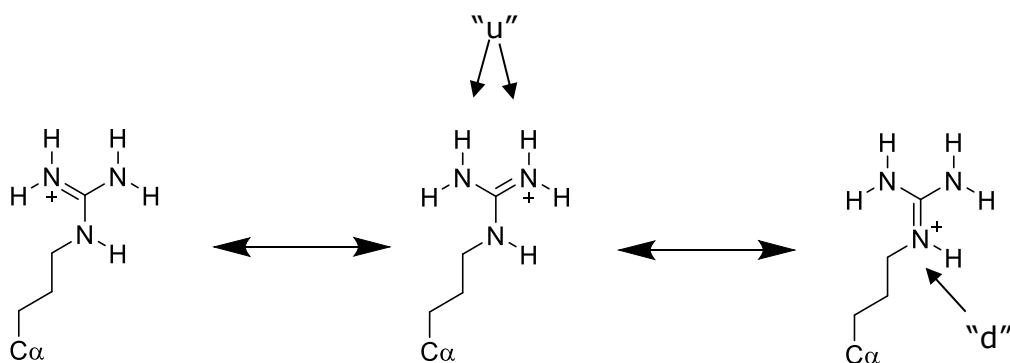


Figure 1-8: Resonance forms of guanidinium ion.

Substitution of a proton for a methyl group causes the removal of a hydrogen bond and as a consequence increases the size and hydrophobicity at that position. This substitution alters the shape of the arginine side chain,^{23,20} however the N-methyl group can still rotate due to resonance.²⁰

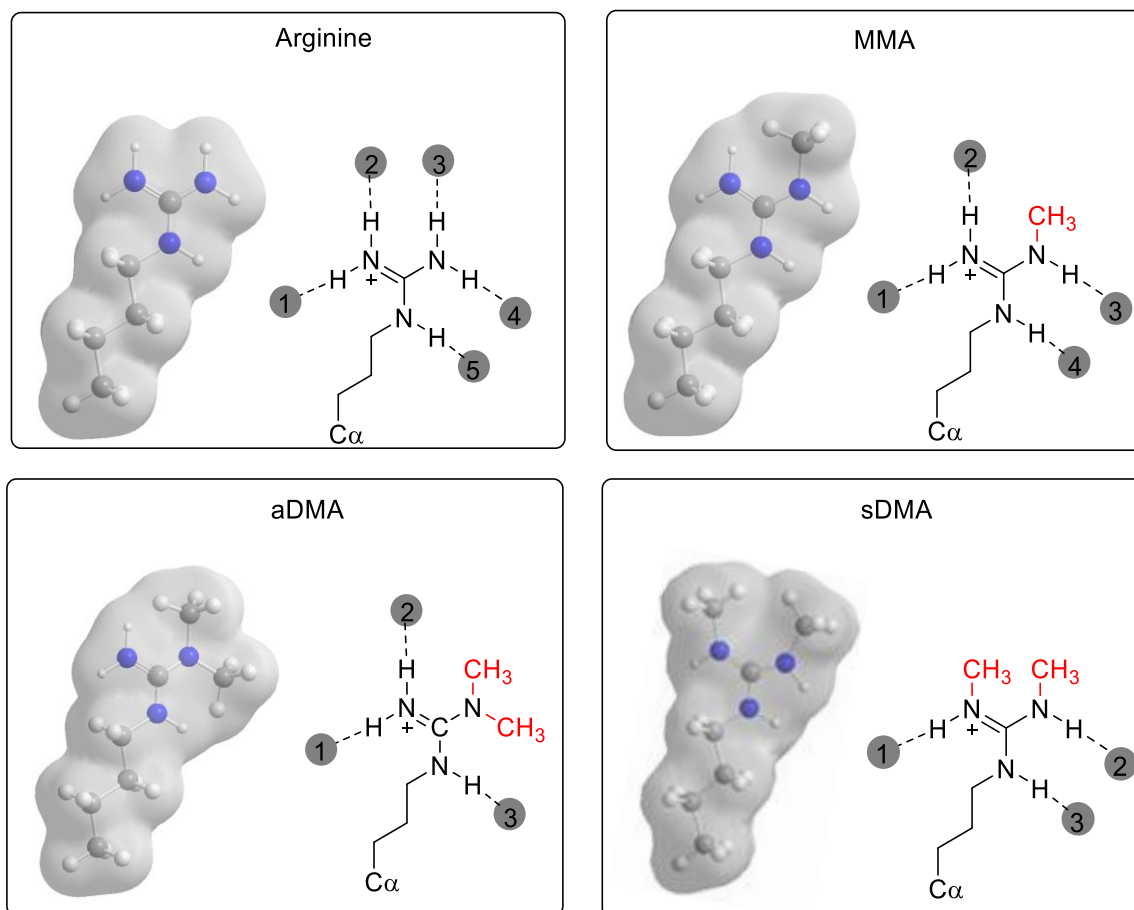


Figure 1-9: The figure shows the surface corresponding to the side chain of the arginine residue. Hydrogen-bonding donor sites and the methylated arginine side chains of MMA, aDMA, and sDMA are shown. The surface was rendered by using ChemBio3D Ultra.²⁰ The stereoisomer of sDMA is omitted for simplicity.

Methylation of the nitrogen does not perturb the positive charge; thus, electrostatic interactions are still available for binding (Figure 1-9).^{20,22} A negligible effect on guanidine pK_{aH} has been observed in experimental studies carried out

with guanidinium salts with various methylation patterns; the pK_{aH} value of guanidinium, 13.6, is slightly reduced to 13.4 upon monomethylation and asymmetrically dimethylated quinidine,²⁰ but is unchanged for symmetrical dimethylguanidine.

1.2.2. The proposed catalytic mechanism

The structure of the catalytic domain of PRMT1 when in complex with SAH and substrate of histone H4 (PDB code 1OR8) was resolved by using X-ray crystallography.²⁴ The crystal CARM1 structures were also resolved when in complex with sinefungin and a histone H3 substrate (PDB code 5DX0).²⁵ Crystallographic studies revealed that SAH and sinefungin analogues bound within the active site of both proteins,²⁵ while substrate arginine residues (H3 for CARM1 or H4 for PRMT1) extend into the substrate arginine binding cavity (Figure 1-10).^{25,24} Notably, the distance between the guanidinium of arginine residue and the primary amine of sinefungin analogues was $\sim 3.0 \text{ \AA}$, which is within range to give optimal interactions for catalysis to occur if sinefungin were replaced by SAM (Figure 1-10).²⁵ Two glutamate residues, Glu-144 and Glu-153 in PRMT1 or Glu-258 and Glu-267 in CARM1, are located at the substrate arginine binding site, which forms electrostatic interactions and a single hydrogen bond to the guanidine group of the arginine residue.^{25,24} These two residues are conserved between PRMTs and play an essential role during methylation. The possible mechanism by which PRMTs transform arginine residue is described below.

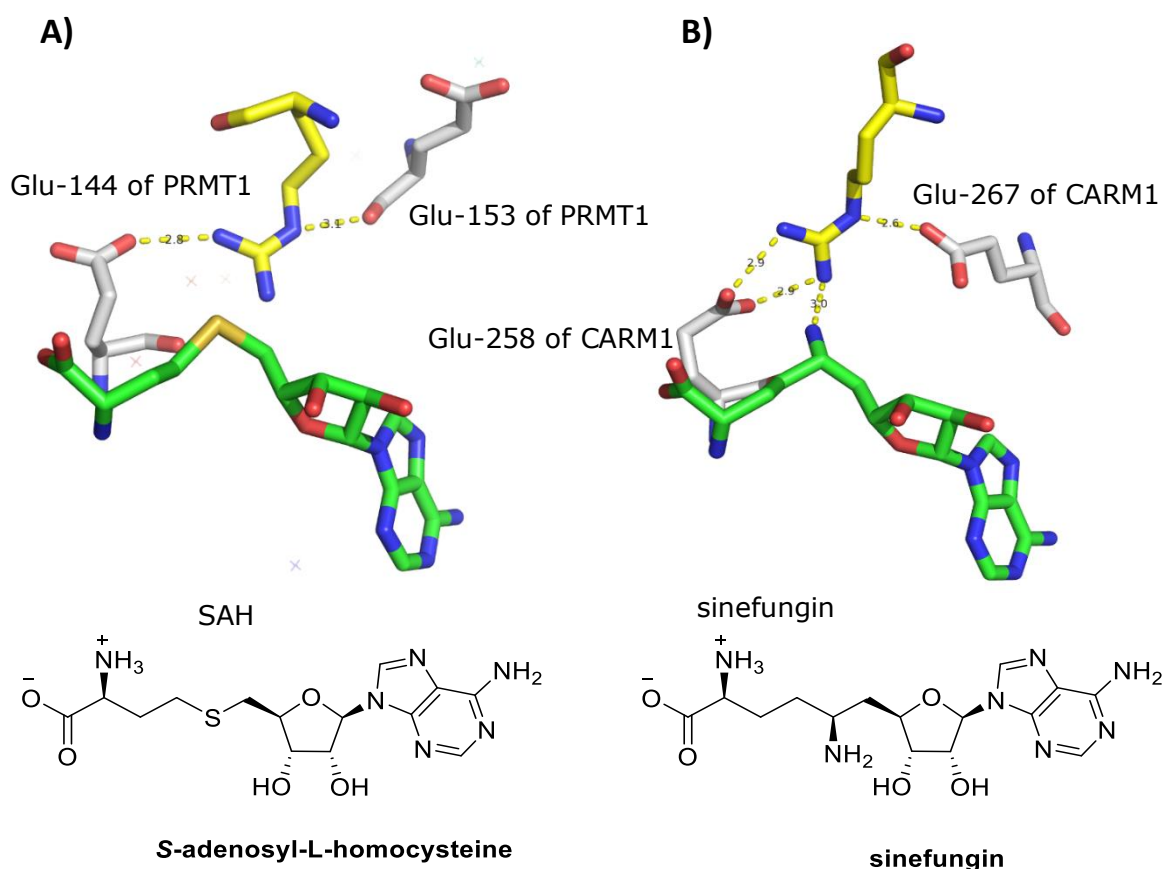


Figure 1-10: Comparison of SAH and sinefungin binding to active site pockets. A) SAH and substrate arginine bound to the active site of PRMT1 (PDB code 1OR8).²⁴ Two glutamate residues form hydrogen bonded with guanidinium of arginine residue are highlighted. B) Two glutamic acid residues Glu-258 and Glu-267 of CARM1 (PDB code 5DX0)²⁵ interact with the arginine substrate.²⁵ Polar contacts of <2.9 Å are represented as dashed lines using PyMol.

Much mechanistic work has been performed on PRMT1 to identify the roles of key residues within the active site,^{18,26} while the following mechanism is proposed for PRMT6. Since a number of residues are highly conserved at the active site of PRMTs (*i.e.*, Tyr-48, His-57, Arg-66, Glu-155, Glu-164, Glu-112 and Met-166, of PRMT6), these findings are generally considered applicable to most PRMTs. These residues are believed to play a key role in recognition of co-factor SAM, peptide substrate, and catalysis.²⁷

Glu-155 and Glu-164 residues make key interactions with the guanidinium functionality of arginine substrate.²⁷ The carboxylate of Glu-164 is believed to modulate the nucleophilicity of the substrate guanidinium *via* charge interactions and two hydrogen bonds with ω -N and δ -N. On the other side, the Glu-155 residue may act together with Arg-66 to orient the atom ω -N' and align the guanidinium ion for nucleophilic attack of the methyl sulfonium ion and leading to a methylated product.¹⁸ It has been assumed that this residue Glu-155 promotes the nucleophilicity of the guanidinium group through hydrogen bonding with deprotonation occurring subsequent to methyl transfer (Figure 1-11). In addition, mutation of either Glu-164 or Glu-155 may cause some structural perturbation and may reduce PRMT6 activity. It has been reported that Glu-144 of PRMT1 when replaced with an aspartate residue and reduces methyltransferase activity by ≥ 3000 times and 10-fold, respectively.²⁴

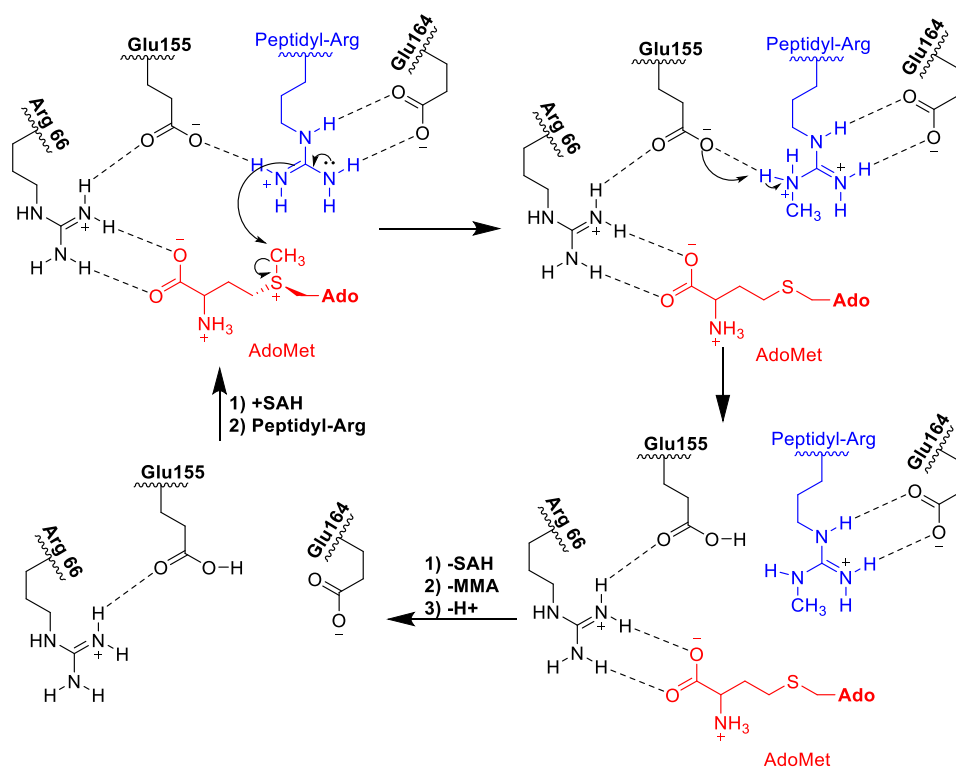


Figure 1-11: The proposed catalytic mechanism for PRMT1 enzymes of type I, exemplified by PRMT6.¹⁸

1.2.3. PRMT substrates and product specificity

Histone tails and heterogeneous nuclear ribonucleoproteins (hnRNP) are prime targets of PRMTs, where methylation of histone arginine plays an essential role in transcriptional regulation.¹⁸ The known histone arginine methylation sites are histone 3 (H3), histone 4 (H4) and histone (H2A), respectively.²⁸ Histone H3 can be methylated by CARM1, PRMT5 and 6, while histone H4 protein can be methylated by PRMT1, 2, 3, 5, 6, 7, 9, respectively. Each enzyme can present selectivity in terms of which arginine is methylated, for example, CARM1 methylates histone H3 at Arg-17, and Arg-26, whereas Arg-8 of the same substrate is methylated by PRMT5 (Figure 1-12).²⁰

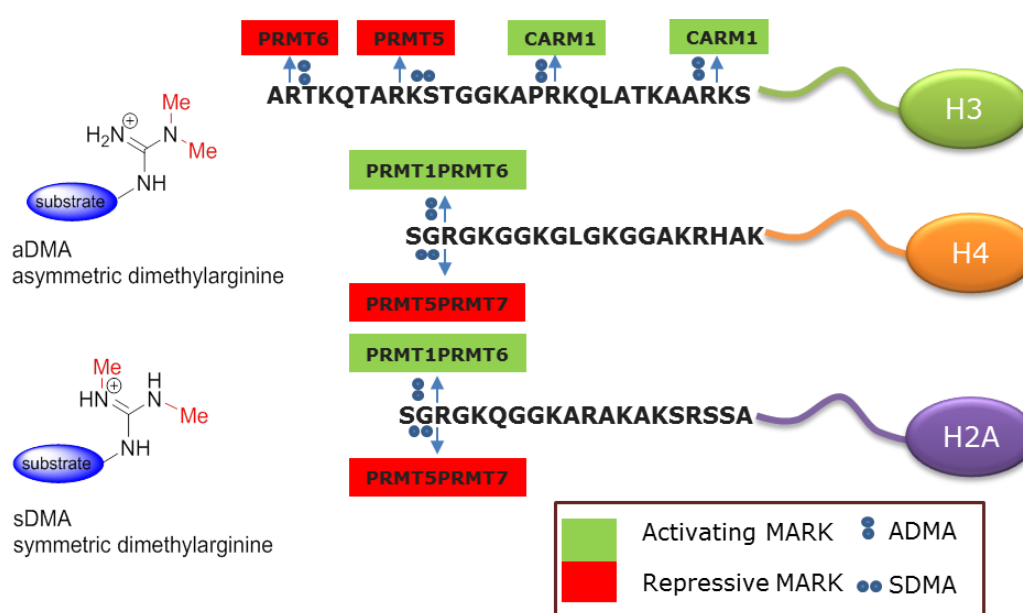


Figure 1-12: Schematic overview of the different methylation sites in histone arginine methylation. Green, activation MARK; red, repression MARK.

Arginine methylation has also been identified in other proteins including nucleolin, fibrillarin and helicases.²⁰ Arginine methylation of nuclear proteins has been found

to be involved in many cellular processes including structural remodelling of chromatin, signal transduction, cellular proliferation, nuclear/cytoplasmic shuttling, translation, gene transcription, DNA repair, RNA processing and mRNA splicing.^{20,5}

A large proportion of protein arginine methylation occurs within glycine/arginine-rich (GAR) motifs,²⁰ except CARM1 which possess a unique sequence proline/glycine/methionine/arginine-rich (PGM) motifs.²⁹

1.3. Structure and functional analysis of PRMTs

In mammalian cells, PRMT proteins can vary from 316 to up to 965 amino acids in size⁷. According to Wolf,⁷ all PRMT proteins possess a common catalytic methyltransferase domain, this domain is a highly conserved region of approximately 310 amino acids; additionally, PRMT1 protein contains subdomains which plays essential roles when binding donor methyl groups to substrate peptides.²⁰ A divergent range of additional protein-domains have been shown to associate with each PRMT; these domains are presumably responsible for directing interactions between enzyme and substrate.

PRMT1 possess an unique catalytic domain region and can methylate arginine residues within glycine/arginine-rich (GAR) motifs.²⁰ PRMT2 contains an SH3 domain which is linked to the catalytic centre; this domain allows PRMT1 protein to interact with proline-rich domains present in other proteins.^{30,31} PRMT3 contains a zinc finger domain,³² and it has been shown that dimethylation of the 40S ribosomal protein, which influences ribosomal biosynthesis, requires the presence

of this zinc-finger domain.^{32,7} CARM1 has an additional Pleckstrin homology (PH) domain at the N-terminus and has a post-core C-terminus extension of around 190 amino acids.^{24,33} PRMT5 has a TIM barrel and has been shown to localise to both the cytoplasm and nucleus.³² PRMT6 is the smallest known protein in the PRMT family, and its role seems to be restricted to the nucleus,³² and whose role seems to be automethylation.³⁴ PRMT7 plays a role in embryonic stem cells (ESC).³² PRMT8 appears to be mainly localised to the brain and has been shown to be attached to the plasma membrane *via* N-terminus myristoylation.³² PRMT9 was identified based on sequence homology to the conserved PRMT SAM-binding motif³⁵ and been shown to localise to both the nucleus and cytoplasm. PRMT9 contains an F-box motif at its N-terminus and a zinc-finger motif at its C-terminus (Figure 1-13).³⁵

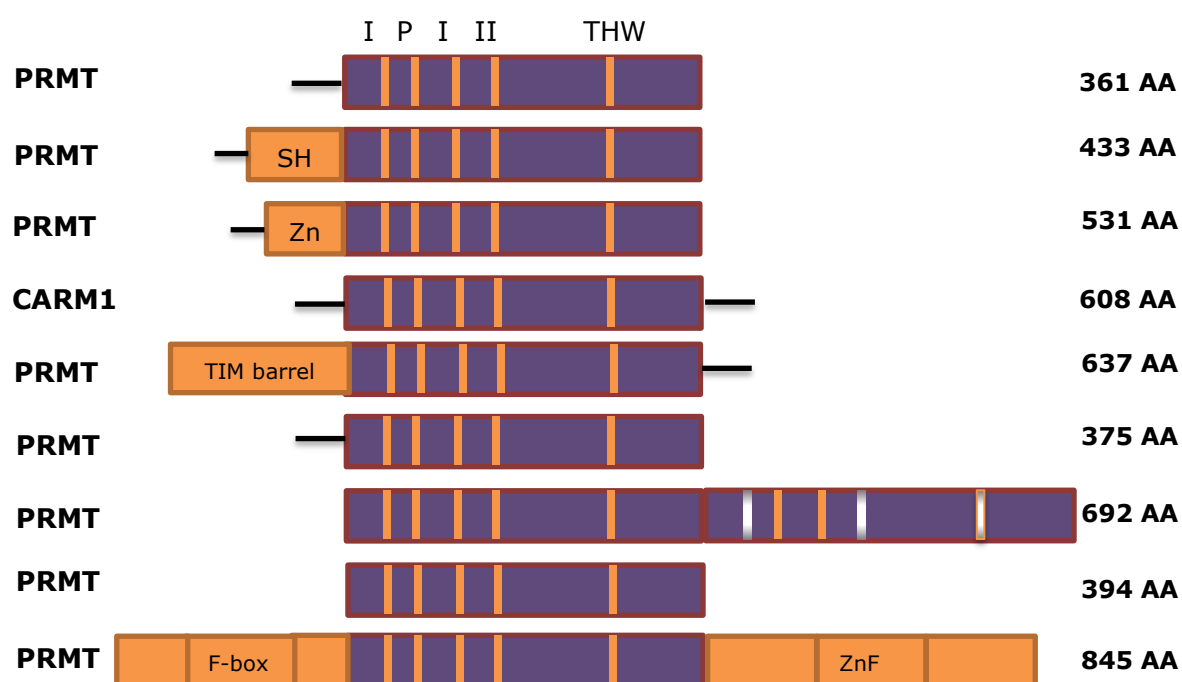


Figure 1-13: Schematic depiction of human PRMT family.

All members of PRMTs family contains signature methyltransferase motifs which include; I, post-I, II, and III and the conserved THW loop, labelled respectively. Some PRMTs have additional domains as shown.²⁰

1.4. Protein arginine methylation and diseases

PRMT1, CARM1 and PRMT6 are the most studied PRMTs, and they have been shown to be associated with a variety of cellular processes such as transcription regulation, alternative splicing, DNA repair, cell cycle regulation and viral immunity.³⁶ A role for PRMTs has been suggested in pathways involved in gene transcription, DNA repair, RNA processing and mRNA splicing.^{34,37} However, dysregulation of these enzymes is associated with several diseases such as cancer, viral diseases and spinal muscular atrophy.³⁷ Increasing prevalence for methylation of arginine residues in substrates was identified in studies with cancer cell lines in the 1970s. Dysregulation of PRMTs can lead to an imbalance of protein levels within cells.^{37,20}

PRMT1 was the first protein arginine methyltransferase discovered in mammalian cells. Dysregulation of PRMT1 has been linked with a variety of diseases, particularly bladder and prostate cancers.²⁰ CARM1 is a coactivator of nuclear receptors, and high expression is correlated with human prostate and breast cancer. PRMT1 and CARM1 overexpression has also been shown in both breast and prostate cancer cell-lines.⁴² PRMT2 has an ability to act as a transcriptional coactivator for ER α and dysregulation of this enzyme has been associated with diseases including breast cancer.²⁰ Additionally, the expression and enzymatic activity of PRMT6 has also been implicated in breast cancer cells.^{20,38}

1.5. PRMT1 and CARM1

PRMT1 is believed to be responsible for approximately 85 % of total enzymatic arginine methylation.³² PRMT1's central role is to regulate protein function, for example, methylation of heterogeneous nuclear ribonucleoprotein (hnRNA) molecules leads to the shuttling of these proteins between the cytoplasm and the nucleus.³² There are no reported human PRMT1 crystal structures; however, a crystal structure of rat PRMT1 (PDB 1OR8) has been published.²⁴ Rat PRMT1 has been shown to have comparable *in vitro* methyltransferase activity with human PRMT1 (Figure 1-14).²⁴

PRMT4 has been shown to methylate histone H3 and non-histone substrates and is more generally viewed as a coactivator-associated arginine methyltransferase (CARM1)³⁹ as it has transcriptional coactivator activity.²⁵ CARM1 has also been shown to associate within the p160 family of proteins.⁴⁰ Numerous crystal structures of CARM1 have been published, many of which emerged during the timespan of this PhD; these include the crystal structure of human CARM1 in complex with *S*-Adenosyl-L-homocysteine (SAH). Three apo crystal structures of CARM1 have been published (PDB: 3B3G, 3B3J, and 2V7E); and several with the enzyme bound to a variety of small molecule inhibitors such as bisubstrate inhibitors and SAH (PDB: 3B3F and 2V74) (Figure 1-14).^{41,42}

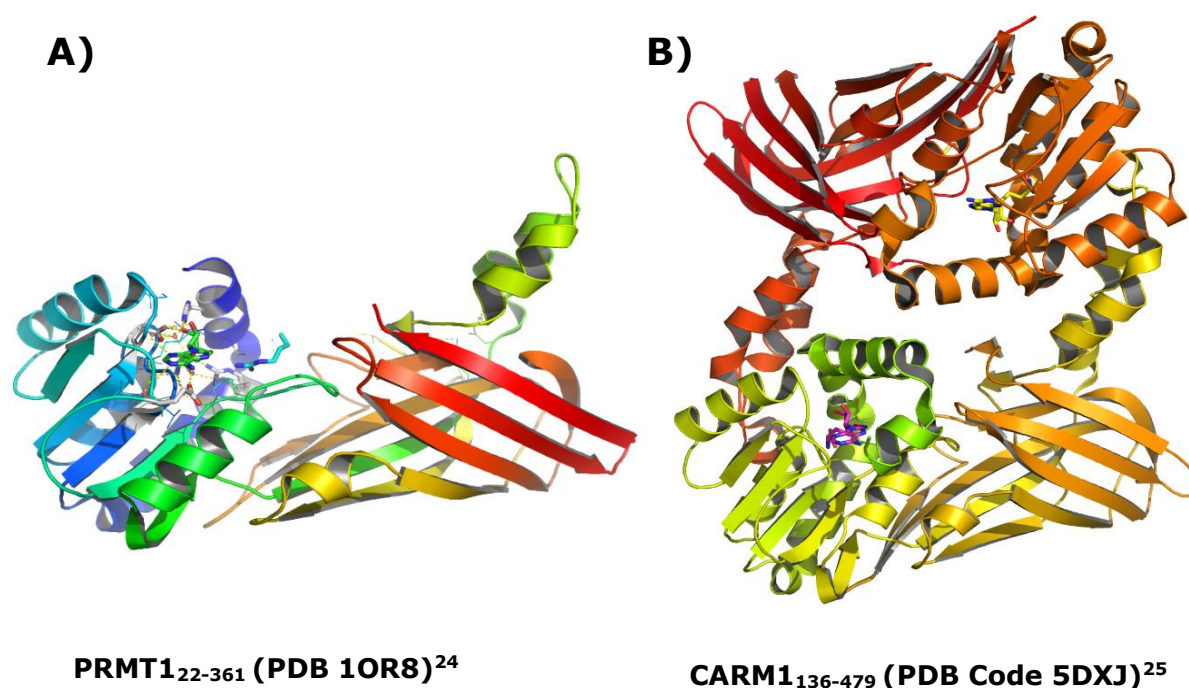


Figure 1-14: Cartoon showed the structures of PRMT1 and CARM1, respectively.
A) The domain structure of PRMT1 in complex with SAH (PDB cod 1OR8),²⁴ B) Dimer structure of CARM1 in complex with SAH (PDB 5DXJ).²⁵

Head-to-tail dimer formation has been observed in crystal structures of PRMT1 and CARM1.^{24,25} It is apparent that this interface is formed by hydrophobic interactions between the dimerisation arm of one PRMT unit with the SAM-binding domain of another.^{24,30,42,25}

1.5.1 Biological roles of PRMT2 and PRMT6

PRMT2 was first identified by homology with the yeast methyltransferase *HMT1* gene.³⁰ PRMT2 is one of the less-characterised members of PRMT family enzymes with regard to structure, activity, and cellular functions.³¹ PRMT2 is a modular protein which contains an SH3 domain which is covalently attached *via* its N-

terminus of the catalytic AdoMet-binding domain.³¹ PRMT2 has been shown to have diverse roles in transcriptional regulation in a variety of cellular processes, through different mechanisms and is dependent on its binding partners.⁴³ Several papers have reported that PRMT2 has weak methyltransferase activity on the arginine residue present in histone H4 substrates.^{31,34} The optimal substrates of PRMT2 have remained unidentified, while PRMT2 methylation patterns remains controversial.³⁴ It is a widely held view that PRMT2 interacts with nuclear hormone receptors, including androgen receptors, retinoic acid receptors, and estrogen receptor alpha, and promotes their transcriptional activity in a ligand-dependent manner.³¹ Besides, its role as a transcriptional coactivator, PRMT2 has also been found to interact with β -catenin, the transcription factor STAT3, the retinoblastoma protein, the early adenoviral region, or 1B-associated protein (E1B-AP5) in a diverse range of cellular processes.^{34,7}

PRMT6 was the first PRMT identified which was capable of automethylation,³⁴ and this function has been reported to be necessary for the maintenance of protein stability.⁴⁴ Enzymatically, PRMT6 has been characterised as a type I enzyme and has been shown to have the ability to methylate arginine residues which are associated with GAR motifs. PRMT6 is considered the primary enzyme responsible for H3R2 methylation in mammalian cells.⁴⁴ PRMT6 may also methylate H4R3/H2AR3 and therefore activate transcription.^{44,32}

1.5.2. PRMT6 structure

Several crystal structures of the full-length PRMT6 have been reported from a number of organisms, including human (hPRMT6 (PDB code 4Y2H)),²⁷ *M. musculus*

(MmPRMT6 (PDB code 5LV5))⁴⁵ and *Trypanosoma brucei* (TbPRMT6 (PDB code 4LWP)).⁴⁶ The hPRMT6 and MmPRMT6 protein sequences shared a 91.5 % homology; however, TbPRMT6 has only around 21-29 % homology when compared to human or mouse PRMT6 proteins. Crystal structures of PRMT6 in the apo form in association with the reaction product S-Adenosyl-L-homocysteine (SAH) and the natural inhibitor sinefungin (SFG)(PDB code 4C03), have been described.³⁴

The catalytic core domain of human PRMT6 consists of around 375 amino acids and contains residues which are highly conserved between PRMTs.²⁷ The catalytic core region of PRMT6 and PRMTs, in general, is composed of three main structurally and functionally distinguishable regions, which include: an N-terminus SAM-binding domain and a C-terminus β -barrel domain interrupted by a helical dimerisation arm domain (Figure 1-15).²⁷ The essential methyltransferase or SAM-binding domain (residues 41-188) is characterised by possessing a unique Rossmann fold and is highly conserved between protein methyltransferases including DNA methyltransferases, histone lysine methyltransferases SET7/9, and indole N-methyltransferase.⁴⁷ The β -barrel domain (residues 190-378) is thought to be unique to PRMTs,²⁴ and it has been shown to be important for substrate binding.⁴⁷ The dimerisation arm is composed of 60 amino acids which are inserted between β 6 and β 7 of the β -barrel domain (residues 200-236).²⁷

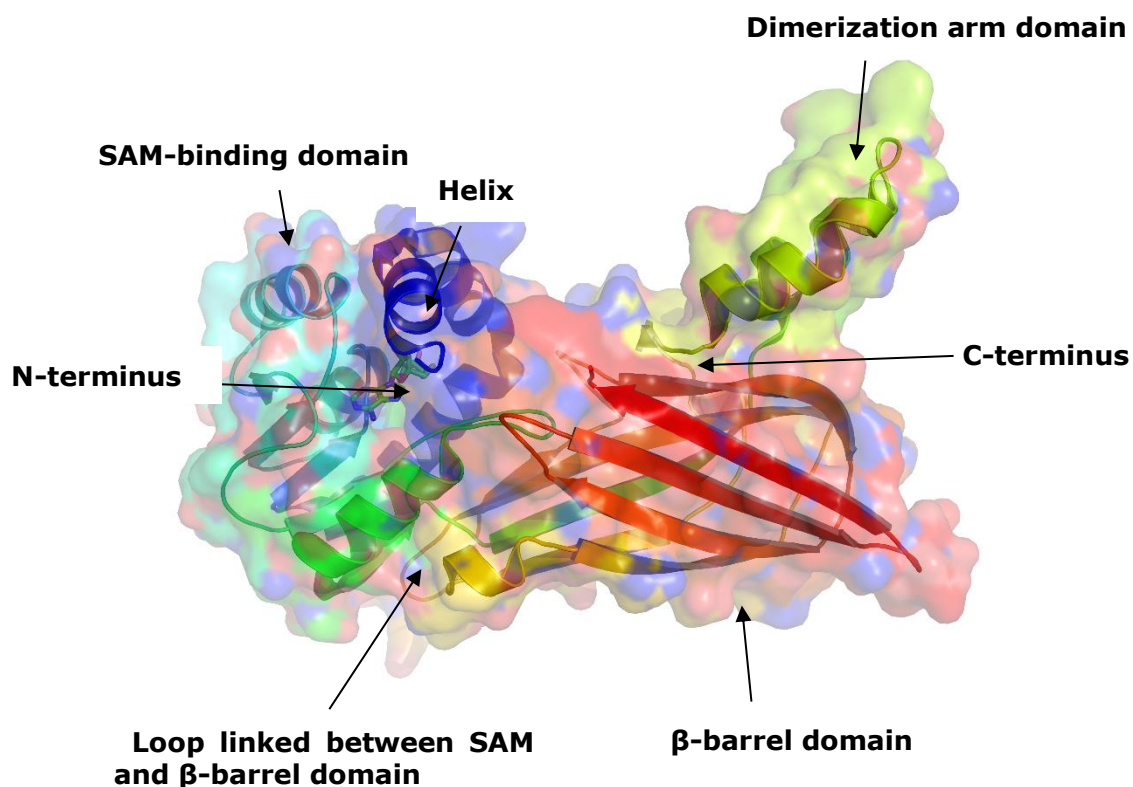


Figure 1-15: The structure of a single monomer of human PRMT6 (PDB code 4Y2H).²⁷

15.3. Homodimerisation of PRMT6

Several PRMT crystal structures have indicated that PRMT6 possesses a ring-like head-to-tail homodimer formation (Figure 1-16);²⁷ this formation is believed to be essential for methylation activity and the formation of the aDMA or sDMA product.²⁰ Generally, PRMTs have been observed to lose activity when the dimerisation arm has been replaced, deleted or mutated.²⁰ The dimerisation arm of one protein unit forms a hydrophobic interaction with the outer surface of a second protein unit at the SAM-binding site (Figure 1-16).²⁷

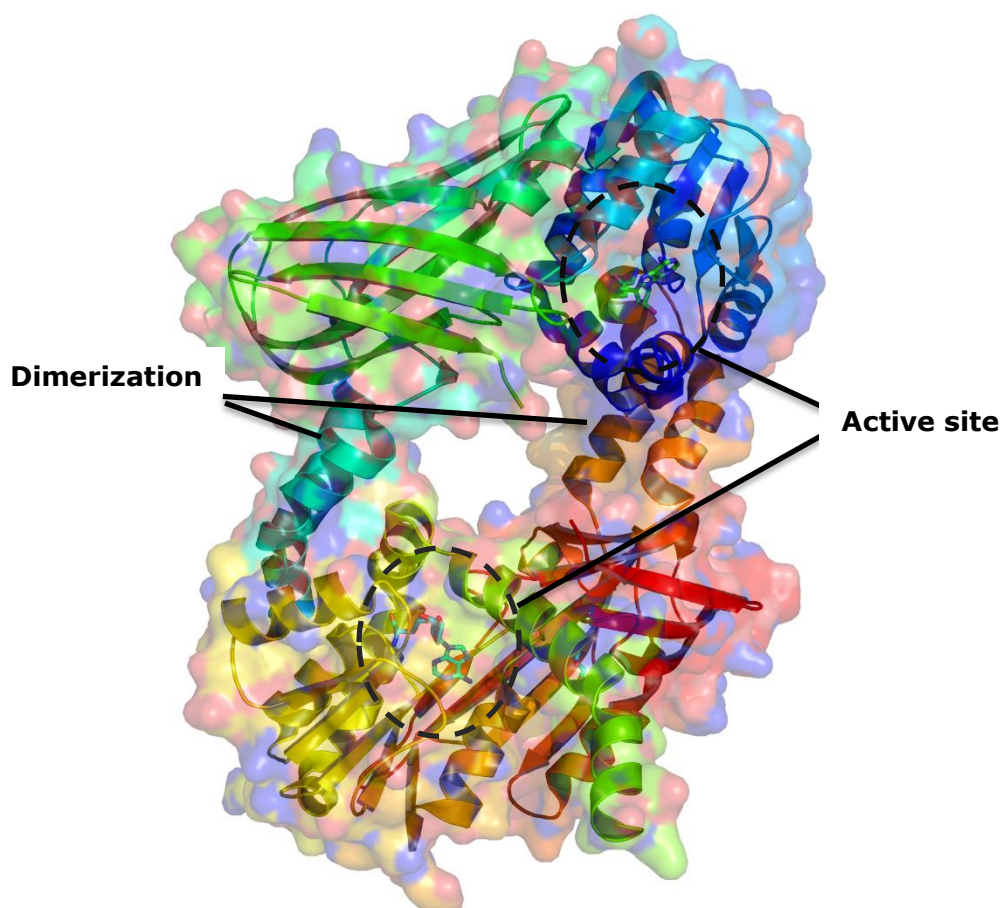


Figure 1-16: Homodimeric PRMT6 contains two active sites and a central hole (PDB code 4Y2H).²⁷

1.5.4. Signature motifs of PRMT6

PRMTs catalytic core region can be split into five main characteristic signature components which includes; motif I, the post I, motif II, motif III and THW loop which surrounds the SAM and target arginine residue binding pockets (Figure 1-17).²⁰

The function of these residues have been described: motif I forms part of the methyltransferase binding site which interacts with the amino acid component of

the co-substrate, this motif is structurally homologous to sequences found in other nucleotide-binding proteins; post-I, which contributes a glutamic acid which is significant for the formation of hydrogen bonding interactions with the two hydroxyl groups of the ribose section of SAM; motif II, which forms a parallel β -sheet which is required to stabilize motif I; motif III, which helps to stabilize the N-terminus helix by forming a parallel β -sheet with motif II; Finally, the THW loop is close to the active site cavity and plays an essential role in substrate recognition.²⁰

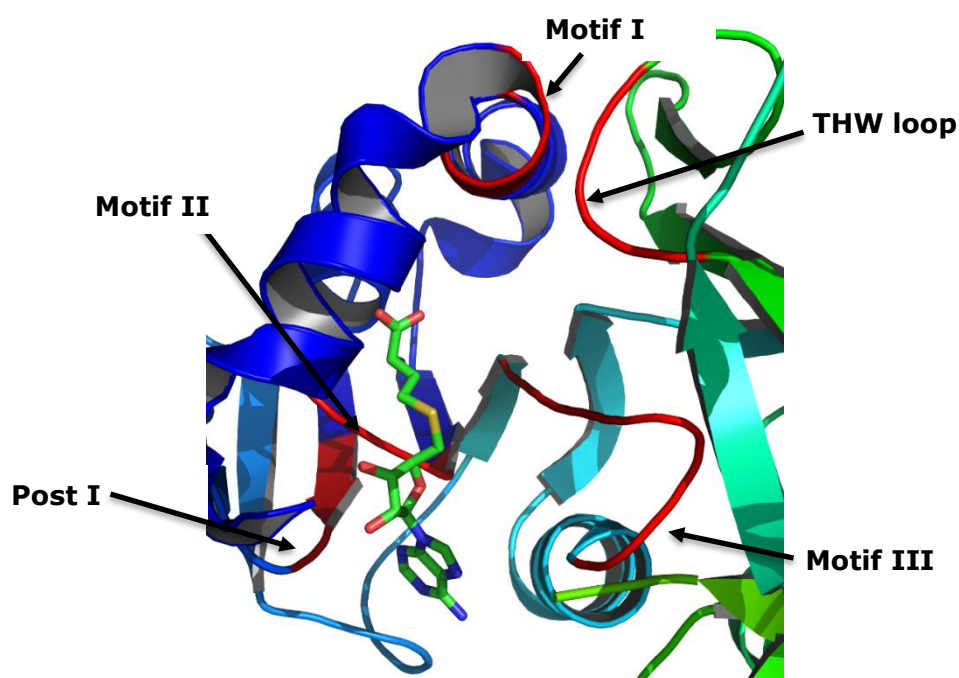


Figure 1-17: Active site of PRMT6 (PDB 4Y2H).²⁷

Overview of the SAH binding site, indicating the PRMT signature motifs I to IV (red backbone).

1.5.5. Co-factor binding mode of PRMT6

The interactions that position the co-product (SAH) and targets arginine residues in PRMT binding pocket are highly conserved.^{20,27} The methionine carboxylate of

SAH interacts with Arg-66 directly and His-57 and Thr-93 *via* a water molecule. The adjacent amine interacts with Asp-88 and backbone protons of Iso-96, Leu-96, Ser-154, and Glu-155 *via* two water molecules in motif 2. The ribosyl hydroxyl groups of SAH interacts with the carboxylate group of Glu-112 in post-I by hydrogen bonding. For the adenine ring moiety, nitrogen N6 interacts with Ser-169, Glu-141 and with Arg-351 *via* a water molecule, while nitrogen N7 interacts directly with Tyr-48 and Arg-351 *via* a water molecule respectively (Figure 1-18).²⁷

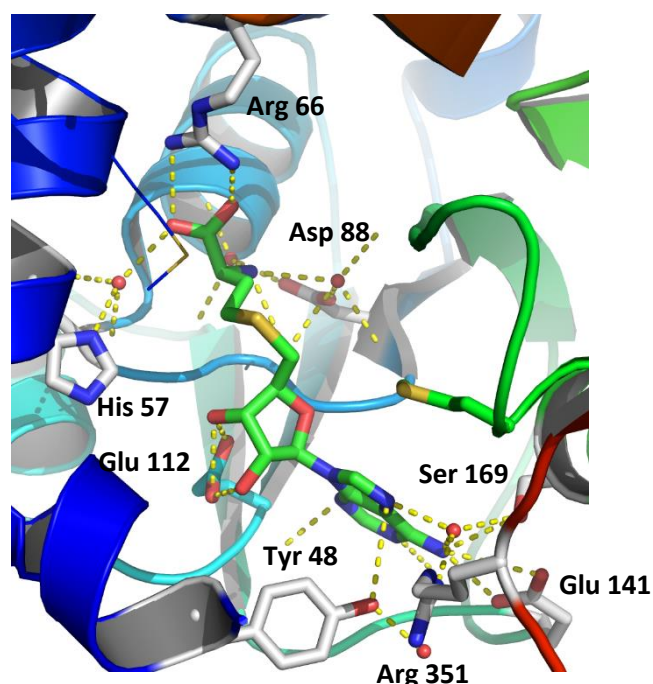


Figure 1-18: Superposition of the catalytic domain of PRMT6 (PDB 4Y2H).²⁷ In the active site, SAH (green backbone) is shown interacting with key residues (white backbone).

The residues lining the arginine peptide substrate-binding channel are not conserved.²⁷ Analysis of a X-ray crystal structure of PRMT6 has revealed that the arginine-binding channel is located in unstructured protein regions with the active site cavity allowing only arginine residues which are present on kinked loops to enter.^{27,20} Two glutamates residues Glu-155 and Glu-164 are conserved in most

PRMTs, and they have been shown to bind to an arginine in the substrate, and this anchors the peptide; this anchoring means the substrate is in contact with other residues, such as Glu-258, Glu-267, Tyr-262 and Met-269 from the N-domain, His-415 from the THWY loop and Asp-166 from αZ , respectively (Figure 1-19).²⁷

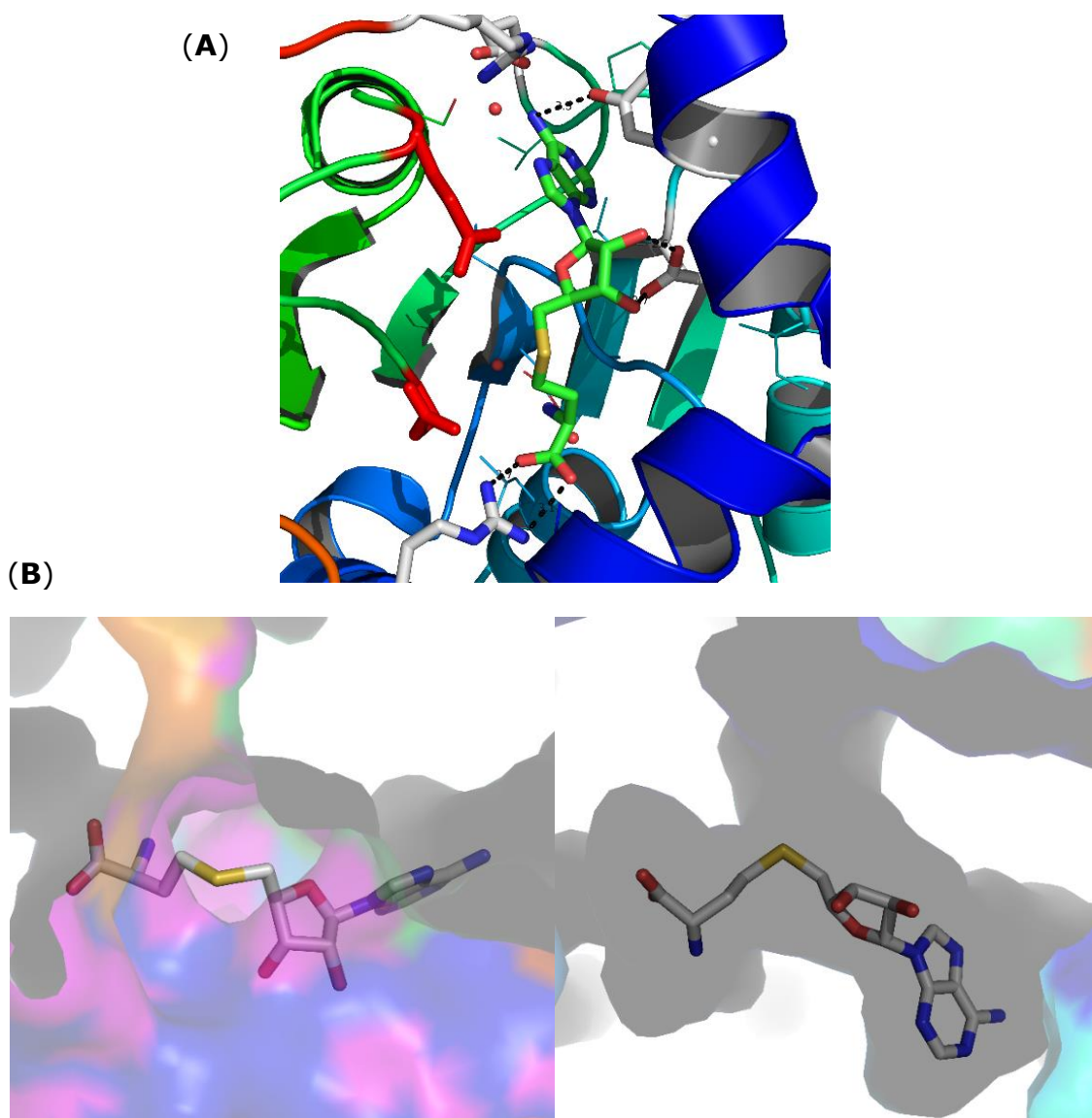


Figure 1-19: Structural of the active site pocket.

A) The binding pocket of the substrate showing conserved two glutamates residues such as Glu-155 and Glu-164, which are highlighted in red. B) The putative substrate arginine channel model. The image represents PRMT6 active site pockets when bound to SAH (PDB 4Y2H).²⁷

Protein lysine methyltransferases (PKMTs, EC 2.1.1.43) have been shown to be capable of methylating one or more methyl groups on lysine residues in histone H3, K-20, K-59 and histone H4, K-26, respectively.⁴⁸ All PKMTs possess a conserved SET domain within the DOTI family, and the co-factor pocket in the active site of these enzymes differs from PRMTs. In contrast, crystallographic studies of the lysine methyltransferase human SET7/9 (PDB 109S)⁴⁹ showed that the peptide substrate and SAH are located on opposite sides of the SET domain. As shown in the Figure 1-20, mono-methylated lysine four on histone H3 and cofactor binding surface are connected *via* a narrow channel; this channel is characterised by the presence of hydrophobic residues which allows passage through the enzyme. Two tyrosine residues (Tyr-245 and Tyr-305) are present in this channel, and they engage in van der Waals interactions with the lysine side chain. The two-tyrosine residue (Tyr-245 and Tyr-305) at the lysine binding site form electrostatic interaction to the lysine ϵ -amino group, aligning it for reaction with SAM.⁵⁰

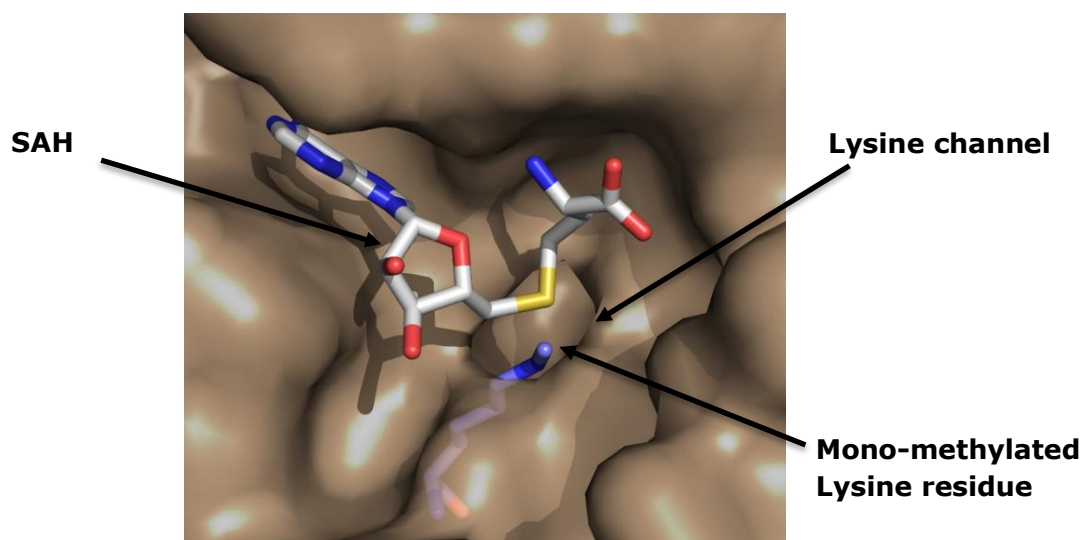


Figure 1-20: Active site pocket of protein lysine methyltransferases.

The active site of the crystal structure of SET7/9 in complex with SAH (cyan) with the substrate lysine (stick) showed as a solid electrostatic surface. There is a narrow channel passing through the enzyme that connects the substrate lysine and SAH binding sites.⁴⁹

1.5.6. Conformations of SAM when in the active site of methyltransferases

Crystallographic studies have revealed that when SAM is bound to the active site of PRMTs, it adopts an extended conformation. Binding of SAH and presumably SAM have both been reported to induce conformational changes within the co-factor pocket in PRMTs producing an access tunnel for target arginine substrates (Figure 1-21).²⁴ Comparison of crystal structures have revealed that SAH binds with a U-shaped conformation to lysine methyltransferases (PKMTs, e.g., human SET7/9 (PDB 109S)),⁴⁹ which is different to the extended conformation observed in PRMTs (e.g., PDB IOR8).²⁴

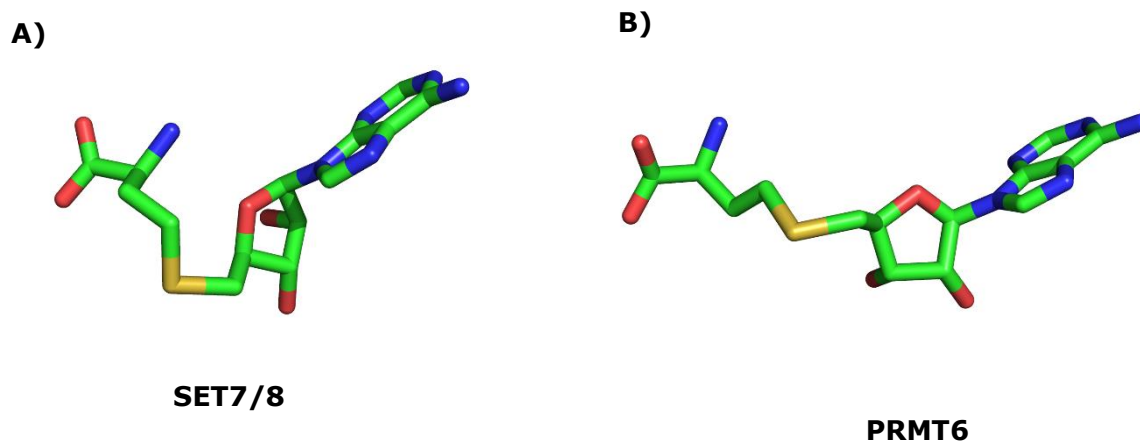


Figure 1-21: Conformation shape of structure SAH showed for comparison.
A) U-shaped conformation of SAH in SET8 (PDB 109S).⁴⁹ B) The extended conformation of SAH in PRMT6 (PDB 4Y2H)²⁷ which is shared in all PRMTs family

1.6. Inhibitors of PRMTs

There is much interest in producing small molecules which affect the cellular activity of PRMTs with the ultimate goal of modulating aberrant methylation processes. Several simple co-factor analogues have been used (Figure 1-22). Adenosine dialdehyde **5** blocks the activity of SAH hydrolase and thus causes intracellular accumulation of SAH **6**, in turn causing indirect global inhibition of PRMTs and other methyltransferase enzymes.^{51,52,53} Methylthioadenosine **7**, occupies the co-factor binding site and displays limited specificity, indiscriminately targeting all SAM-dependent enzyme with typical IC_{50} values of 0.1 - 20 μM .^{53,54} Sinefungin **8** is a natural product that was isolated originally from *Streptomyces griseus* and is a potent but indiscriminate inhibitor of all SAM utilising enzymes, due to its structural similarity with the co-factor.⁵¹ A chiral amine, which is protonated at physiological pH, serves as a stable mimic of the sulfonium species displayed by the co-factor. A common feature of these molecules is that they are indiscriminate, affecting the action of all SAM utilising enzymes.²⁰

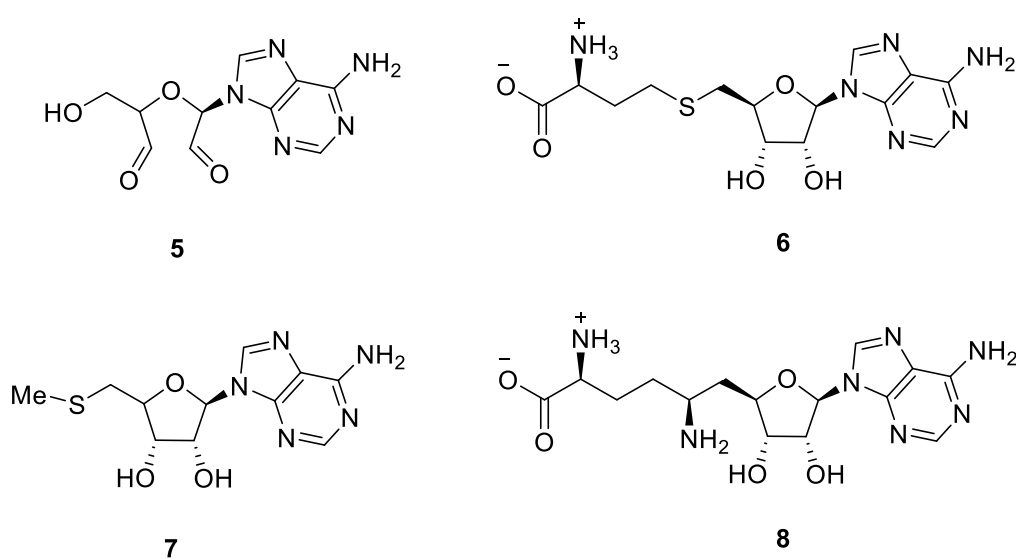


Figure 1-22: A selection of nonselective PRMT inhibitors.

A greater variety of PRMTs inhibitors have emerged over the last fifteen years. A small-molecule inhibitor of PRMTs discovered by high throughput screening was first reported by Cheng in 2004.⁵⁵ The symmetrical sulfonated urea AMI **9** (arginine methyltransferase inhibitor) was first reported to selectively inhibit arginine methyltransferases over lysine methyltransferases with an IC₅₀ of 8.8 μM.⁵⁵ Moreover, compounds **10** and **11** were shown to selectively inhibit PRMT1 with IC₅₀ values of 41.7 μM and 12.0 μM, respectively.⁵⁶ Inspired by this compound's structure, compound **12** was recently reported to inhibit PRMT3 selectively and was confirmed by mechanism of action (MOA) studies to be non-competitive with both the peptide substrate and the cofactor.^{56,57} It has been reported that PRMT3 has an allosteric site which utilised binding pockets other than the cofactor binding site (Figure 1-23).⁵⁷

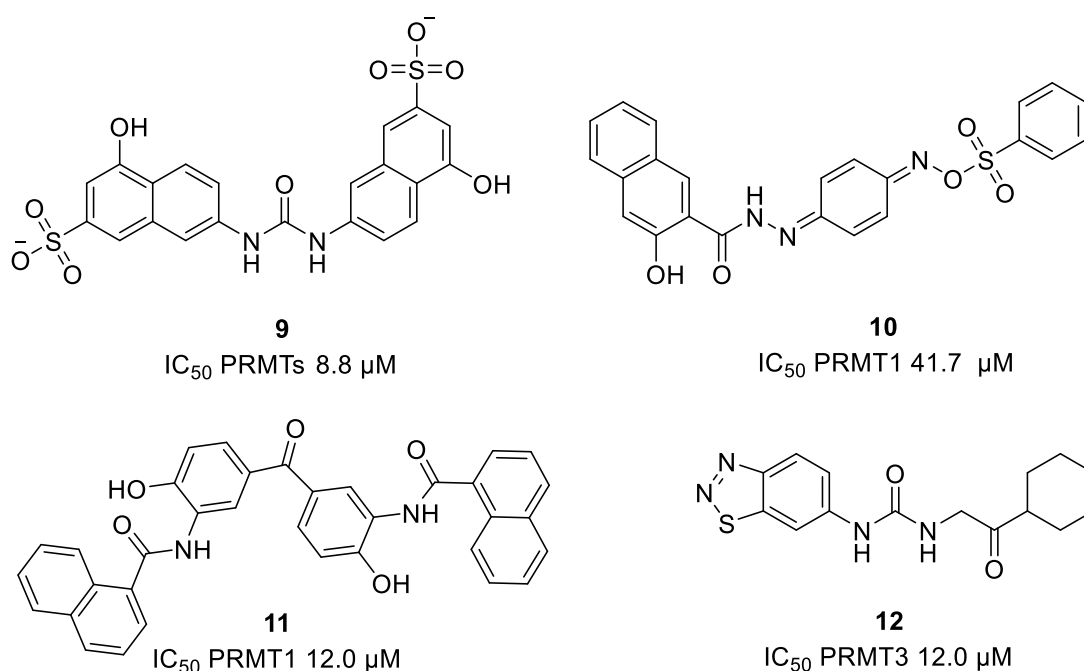


Figure 1-23: A selection of molecules which have been shown to act as PRMT inhibitors.

Recently, a number of more drug-like molecules have been developed as PRMT inhibitors. For example, inhibitor **13** has emerged as a potent inhibitor of CARM1 with an $IC_{50} = 6.0 \pm 3.0$ nM; this compound also displays selectivity over other histone methyltransferases.⁴⁰ More recently, compound **14** has been reported as a PRMT6 inhibitor with an IC_{50} of 10.0 nM.^{58,27} Further development and optimisation have identified more potent CARM1 inhibitors such as indole **15** and pyrazole **16**.⁴¹ These inhibitors have IC_{50} values against CARM1 of between 30.0 ± 15.0 nM and 27.0 ± 9.0 nM respectively; they also display a weak inhibitory effect against PRMT1 and PRMT3 with IC_{50} values of >10 μ M.^{27,41}

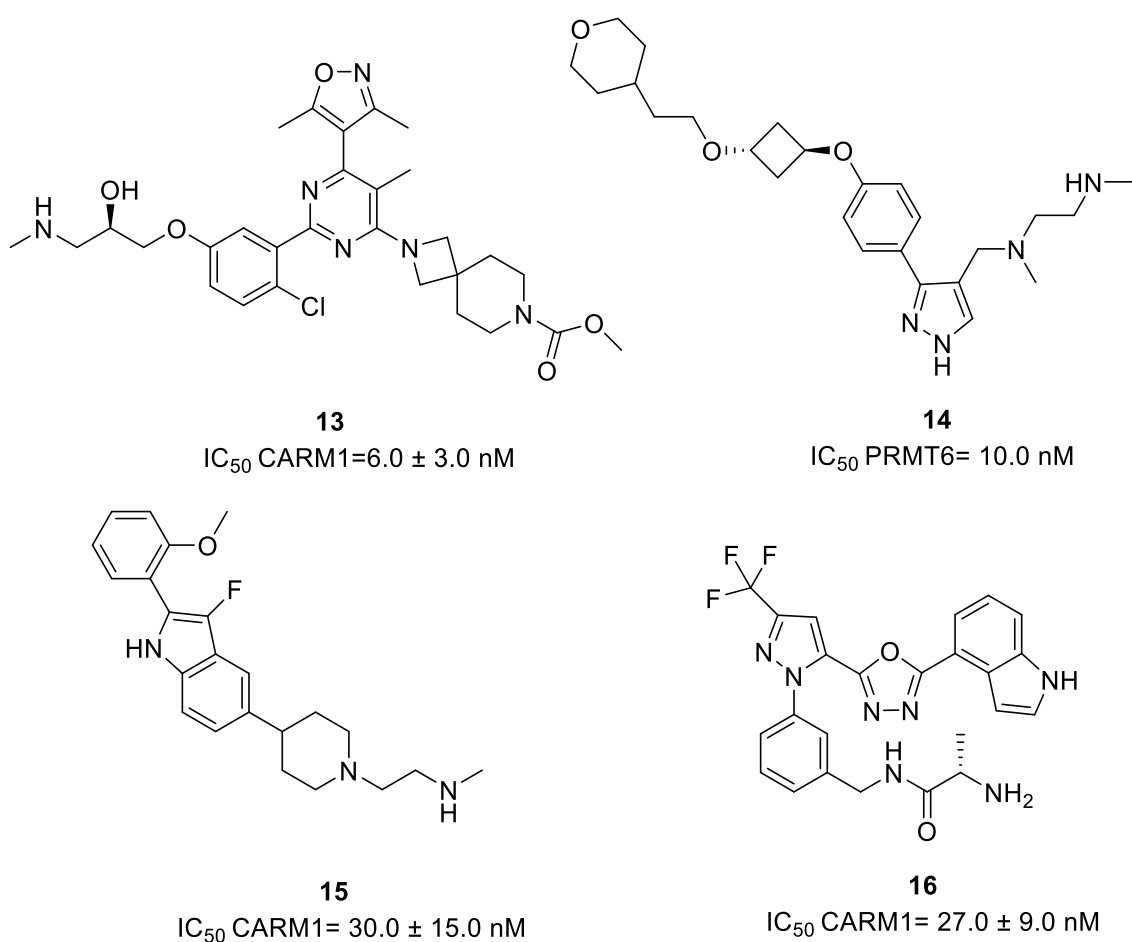
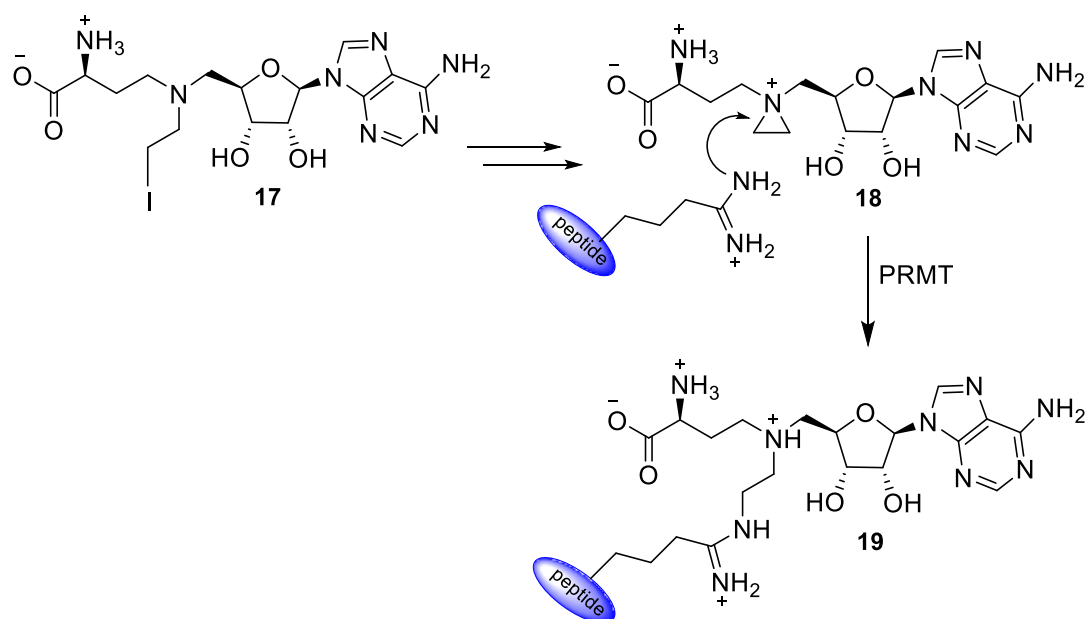


Figure 1-24: Chemical structure of some PRMTs inhibitors which have been reported as being inhibitors of CARM1 and PRMT6, respectively.

1.6.1. Bisubstrate inhibitors

Bisubstrate inhibitors are an effective strategy for facilitating and promoting the potency and selectivity of PRMTs over other enzyme inhibition.⁵⁹ It has been suggested that covalently joining mimics of the two co-substrates may provide a high degree of specificity for these enzymes.^{16,20} This result may be explained by the fact that bisubstrate species may not be recognised by other enzymes that utilise the individual components. Thermodynamically, if this bisubstrate reaches a critical size, the binding of a bisubstrate inhibitor to the pocket of these enzymes becomes energetically favourable when compared to the binding of the two substrates separately. The past decade has seen a rapid development of bisubstrate inhibitors for many human methyltransferases such as histone methyltransferases SET7/9, spermine synthase, catechol O-methyltransferase, DNA methyltransferases and indole *N*-methyltransferase.^{20,16}

Inhibitor compound **19** has been reported to inhibit PRMT1 activity with an IC₅₀ of 18.5 μ M and has a four-fold selectivity against PRMT1 when compared with CARM1.⁶⁰ It is thought that the aziridinium moiety of 5'-N-iodoethylamino-5'-deoxyadenosine **17** is susceptible to nucleophile attack of the peptide forming highly reactive and unstable aziridinoadenosine **18**.⁶⁰ Interestingly, upon incubation with PRMT1 and substrate of H4, this bisubstrate inhibitor alkylated the arginine residue of a peptide (H4R3) in a reaction catalysed by PRMT1 to form compound **19** (Scheme 1).⁶⁰ Despite the promising *in vitro* IC₅₀ values reported for these peptide-based bisubstrate analogues, their use has been associated with undesirable pharmacological properties.⁶⁰



Scheme 1. Reactions catalysed by PRMT1 and modification of an arginine residue. SAM derivative 5'-N-iodoethylamino-5'-deoxyadenosine is transformed *in situ*, which when reacted in the presence of PRMT1 generates a bisubstrate PRMT inhibitor.⁶⁰ The blue sphere represents a peptidyl arginine substrate.

Several bisubstrate inhibitors which incorporate a peptide substrate and the α -amino carboxylic acid moiety of SAH have also been reported as PRMT inhibitors. In 2011, SAM derivatives conjugated to a guanidinium group *via* varying carbon linkers (Figure 1-25) were reported.⁶¹ An analogue which contains a guanidine *via* a three-carbon linker **22** inhibited PRMT1 ($IC_{50} = 6.2 \pm 3.8 \mu M$) and CARM1 ($IC_{50} = 13.3 \pm 8.7 \mu M$) without inhibiting lysine methyltransferase ($IC_{50} = >100 \mu M$).⁶¹ Remarkably, the compound that possessed a guanidine group *via* a longer carbon chain (four **23** or five **24**-atom long) were generally more active against PRMT1 and displayed little or no inhibition of CARM1 or lysine methyltransferases, respectively. More rigid inhibitors which contained a benzylic linker **20** inhibited PRMT1 with an IC_{50} value of $5.6 \pm 4.7 \mu M$ but not CARM1 or lysine

methyltransferase SET7, respectively.⁵⁹ Recently reported, a new partial bisubstrate PRMT inhibitor **21** displays potency towards PRMT1, PRMT4, and PRMT6 with IC₅₀ values of 1.3 μ M, 0.56 μ M, and 0.72 μ M, respectively.⁶²

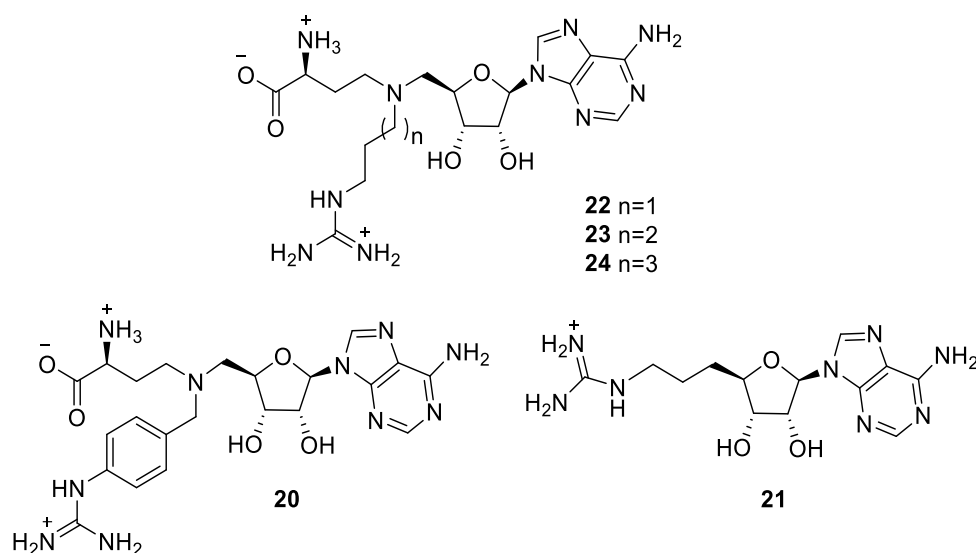


Figure 1-25: Recent developments in chemical structures of PRMTs inhibitors.

The Dowden group has investigated the binding of compounds **20**, **22-24** by *in silico* docking⁶³ of compounds against the crystal structure of CARM1 (PDB 2Y1X)⁴⁰ using Autodock Vina.^{61,63} AzaSAM analogue **22** was found to bind at both the SAM and arginine substrate binding sites. Analysis revealed that the guanidine functionality of compound **22** extended into the substrate arginine binding cavity and interacted with key residues Glu258 and Glu267²⁵ (Figure 1-26).

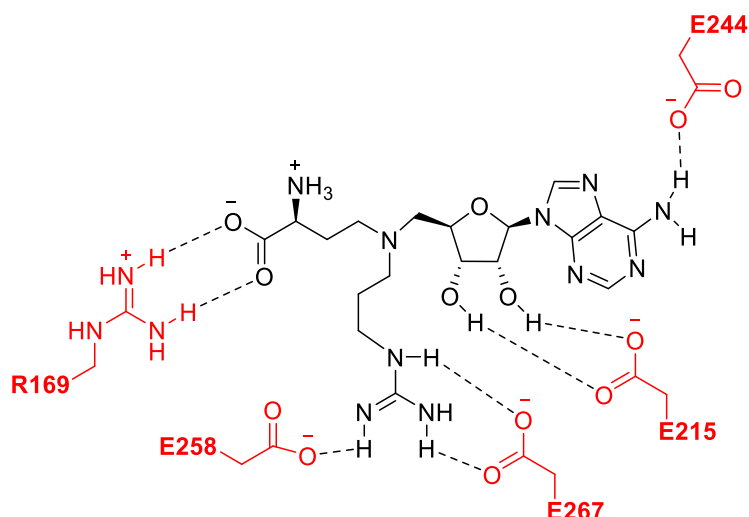
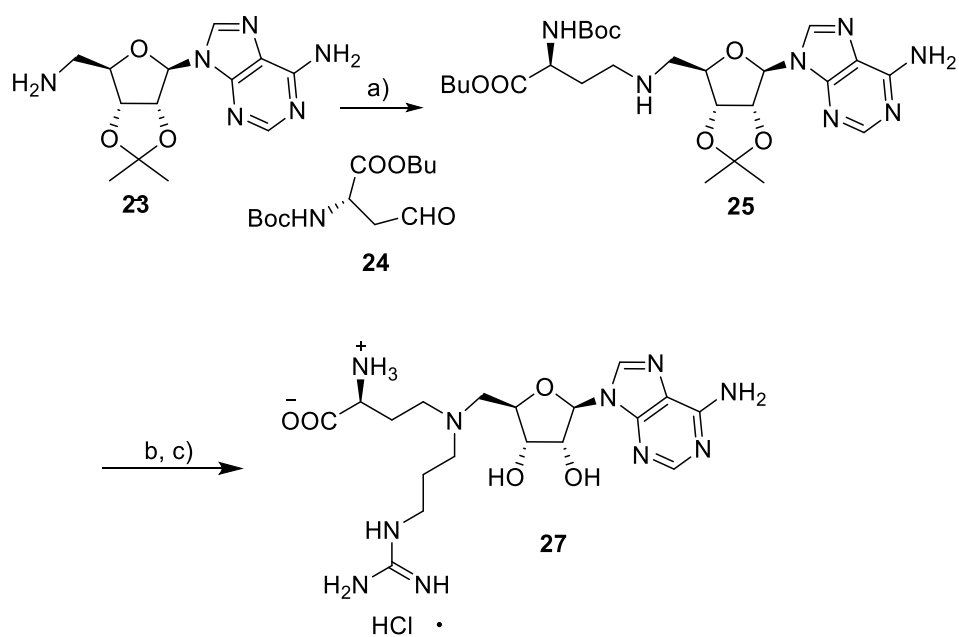


Figure 1-26: General concept for the design of AzaSAM analogue.

Key predicted interactions between the residues of CARM1²⁵ within SAM and substrate arginine binding sites of bisubstrate inhibitors.

There is a synthetic route for the modular installation of readily accessible aldehyde precursors and this route relies on sequential reductive amination steps (Scheme 1-2). Secondary amine **25** was synthesised by initial alkylation of derived aldehyde protected **24** with amine 5'-amino-5'-deoxyadenosine **23** using NaBH(OAc)₃ to produce compound **25**, with a 73 % yield. Subsequent reductive amination with derived aldehyde protected in the presence of NaBH(OAc)₃ generated the bisubstrate inhibitors **27** after deprotection with neat TFA and counterion exchange with a 90 % (v/v) overall yield (Scheme 2).^{3,2}

Using an *in vitro* methylation assay, compound **27** inhibited PRMT1 (IC₅₀ = 6.2 ± 3.8 μM) and had no significant activity against protein lysine methyltransferase SET domain.^{2,3} These promising results encouraged us to seek more potent and selective inhibitors against individual PRMT members.



Scheme 2: Synthesis of AzaSAM analogue.

Reagents and conditions: a) NaBH(OAc)_3 , $\text{ClCH}_2\text{CH}_2\text{Cl}$, r.t., 2h, 73 %; b) NaBH(OAc)_3 , $\text{ClCH}_2\text{CH}_2\text{Cl}$, r.t., 2h, 60 %; c) 9/1 THF/ H_2O , r.t., 4h, 67-90 %.

During work towards this thesis, Cura, *et al.*, (2017) disclosed a new bisubstrate inhibitor that co-crystallised with CARM1. Structures have been resolved for several of these compounds including bisubstrate inhibitors **28** (PDB code 5IS8), **29** (PDB code 5ISA), **30** (PDB code 5IS9) and **31** (PDB code 5LKJ) (Figure 1-27). At the time of initiating this project, there were no crystal structures pertaining to the inhibitors developed in the Dowden group.

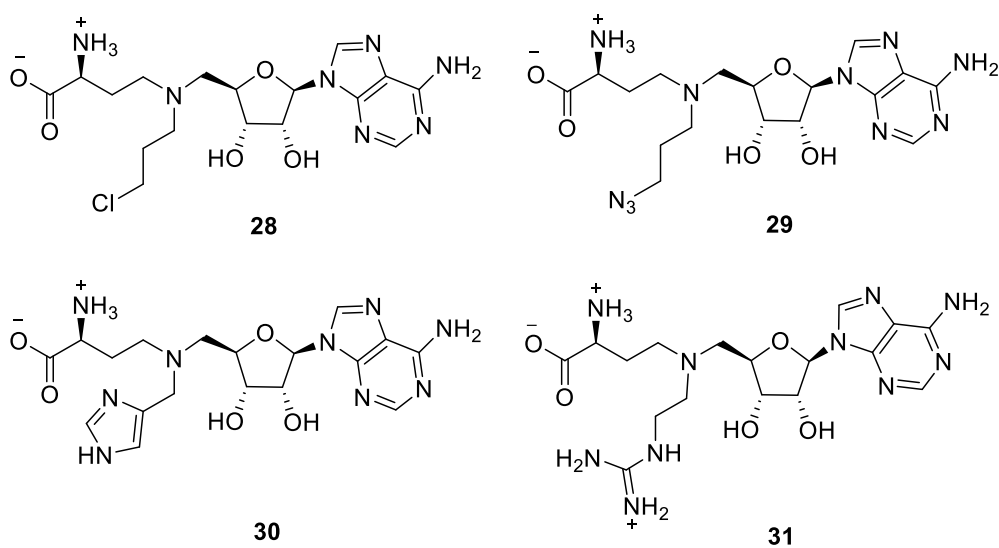


Figure 1-27: Recent development in the chemical structure of PRMTs inhibitors.

1.7. Project aims

The aim of this thesis was a combination of the development of protein expression/purification of selected PRMTs and the design and synthesis of novel chemical probes. The first objective was to generate small molecule inhibitors which were potent and selective for individual PRMTs focusing on target enzymes, PRMT1 and CARM1. These inhibitors were developed as AzaSAM analogues, which combined the structural features of SAM and a substrate arginine motif. Specifically, we wanted to investigate the effect of substituting the guanidinium groups with isosteric replacements.³ The design and synthesis of mimic compounds were focused on substituting the guanidinium group with aminopyridine with the aim of decreasing overall inhibitor polarity.

The second objective was to achieve a biological evaluation of these molecules toward PRMT1, CARM1 and PRMT6 enzymes, and to characterise the binding of these small molecule inhibitors using isothermal titration calorimetry and an *in vitro* methylation assay.

The third objective was to co-crystallise the bisubstrate inhibitor with target enzymes: PRMT1 and CARM1, providing valuable structural information of the inhibitor's position. Additional objectives were to investigate the structure of PRMT2 by crystallography and determine how the substrate and the inhibitor bound to this enzyme.

1.8. Prototype PRMT inhibitors

In the present study, the focus was to systematically vary functional group substitution and evaluate their effect on inhibition of PRMTs. A particular interest was in replacing the highly basic guanidinium group with more hydrophobic replacements which did not carry a permanent charge. This would be an important step towards making such a molecule cell-permeant, a requirement for utility as potential chemical probes of cellular PRMT activity. Selective small-molecule inhibitors of PRMTs are potentially important tools for investigating the biology of this emerging target.² The same basic design was followed for the development of novel bisubstrate inhibitors. Therefore, AzaSAM analogues **55** and **58** were designed and synthesised (Figure 1-28), which contained pyridine and terminal propylamine groups, and terminal α -amino acid group as guanidine isosteres.

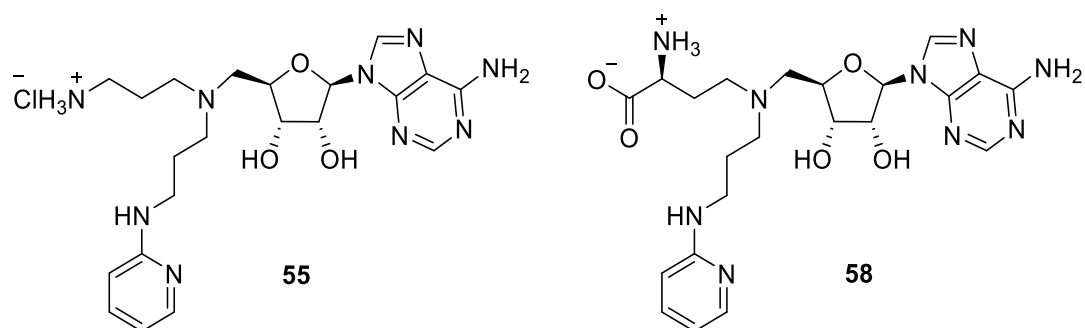


Figure 1-28: Chemical structure of AzaAdoMet analogues.

Chemical structures of target bisubstrate inhibitors are containing different linker groups, such as pyridine and terminal propylamine groups and a terminal α -amino acid group.

These promising results encouraged us to generate additional inhibitors towards PRMTs. AzaSAM analogues with various carbon linkers, such as compounds

contain propylamine groups, and terminal α -amino acid group **46** were also generated as novel selective probes for PRMTs. Lactam **39** was generated by cyclisation of the secondary amine with an aspartic acid derived aldehyde bearing benzyl ester; discussed further in chapter two. Furthermore, a compound containing guanidinium group **64** was also synthesised,⁶² with the aim of comparing its binding to our bisubstrate inhibitors with other protein arginine methyltransferases being prepared during the project. This compound was reported by the Martin group and reported to be more active towards CARM1 with an IC_{50} $0.12 \pm 0.02 \mu M$, and to be a slightly weaker inhibitor of both PRMT1 and PRMT6 with an IC_{50} 11.09 ± 2.77 and $20.23 \pm 8.67 \mu M$, respectively (Figure 1-29).⁶²

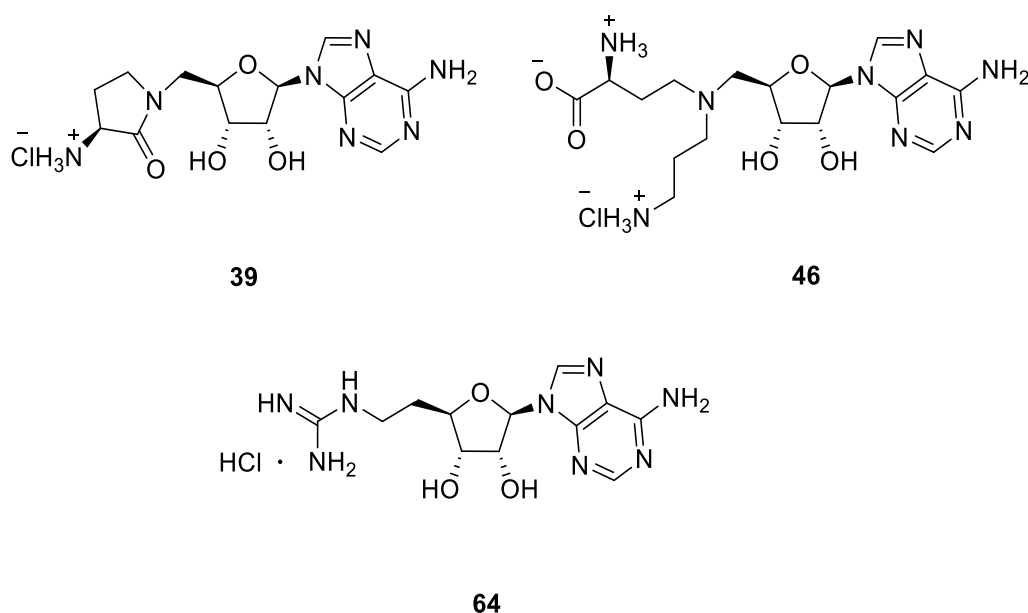


Figure 1-29: Chemical structures of AzaSAM analogues containing different linker carbons.

1.9. Preparation of hydrophobic version for SAM analogue

Synthesis of a hydrophobic version of SAM analogues **74** and **87** (Figure 1-30) was attempted with the aim of exploring whether these compounds could cross cell membranes.^{64,65} Based on this investigation, the production of a hydrophobic version of the sulfur-linked compound by adding butyl esters at the 2',3' hydroxyl was focused on. If successful, this platform was used to deliver co-factor analogues, such as sulfonium species which has a functionality separate from that observed with a methyl group. A hydrophobic version of SAM analogues was synthesised with different side chains, such as propylamine **74**, and α -amino carboxylic acid moiety of SAH **87** (Figure 1-30). It is possible to hypothesise that the butyl ester of these compounds may hydrolyse within the cell to deliver SAM or an analogue which might compete with SAM at the enzyme binding site.

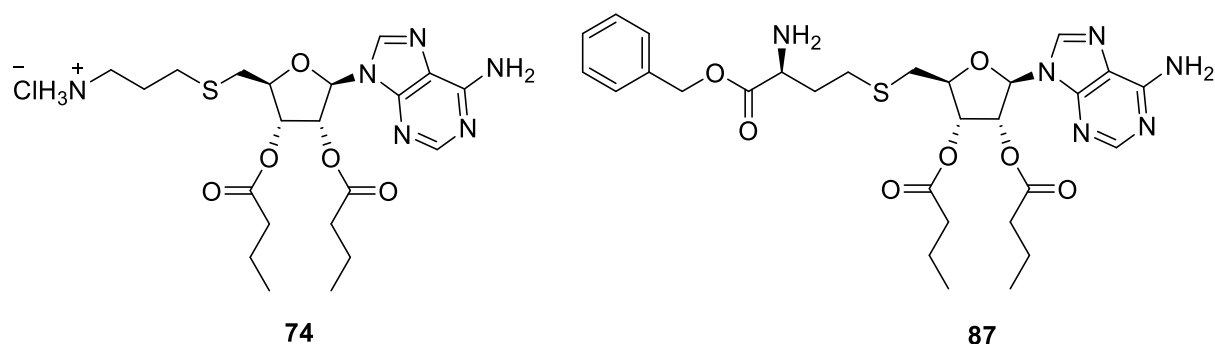


Figure 1-30: Chemical structures of SAM analogues.

SAM analogues with the hydrophobic version of butyl esters at the 2',3' hydroxyl group.

Chapter 2 Chemistry results and discussion

2.1. Aims and objectives

The main objective of this part of the project was to build upon findings from previous group members and to prepare bisubstrate inhibitors designed to include the AzaSAM structure, as well as that of the peptidyl arginine substrate.^{2,3} Previously, related inhibitors featuring a guanidinium group linked to AzaSAM *via* an alkyl chain containing three, four or five methylene groups were synthesised and were found to be PRMT1 inhibitors.⁶¹ Surprisingly, inhibitory compounds which contained four and five methylene-linked groups displayed little or no inhibition of CARM1. New inhibitors included compounds, where the guanidinium group was replaced by another functional group, were synthesised; these inhibitors maintained the hydrogen-bonding pattern but were less basic and were not positively charged at physiological pH (Figure 2-1). Based on modelling, it was expected that these molecules would interact with key residues of the active site in conserved PRMTs, such as Glu-144 and Glu-153 of CARM1.²⁵

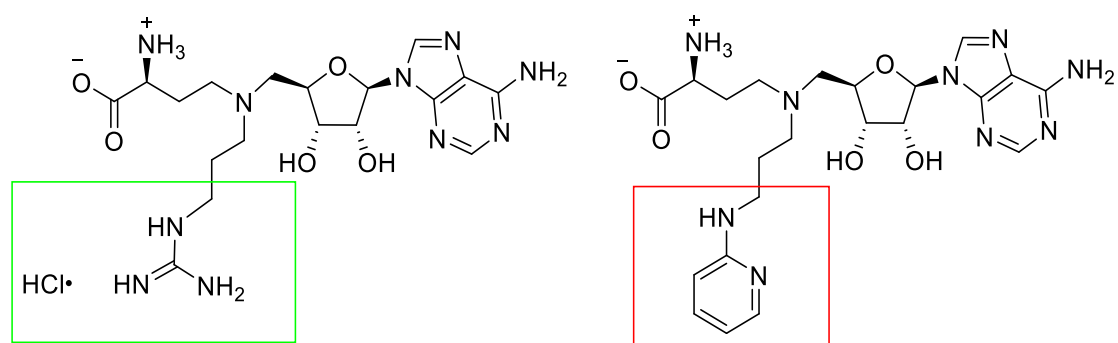


Figure 2-1: Structure of bisubstrate inhibitors.

Guanidinium group highlighted with green replaced by another functional group pyridine, highlighted in red.

A synthetic route to make a version of SAM that incorporated *n*-butyl esters, at the 2' and 3' hydroxyls of adenosine was also developed (Figure 2-2). The idea here is to later develop co-factor analogues that may be able to pass through the cell membrane. Progress towards this objective will be discussed in more detail later in this chapter (see Section 2.2).

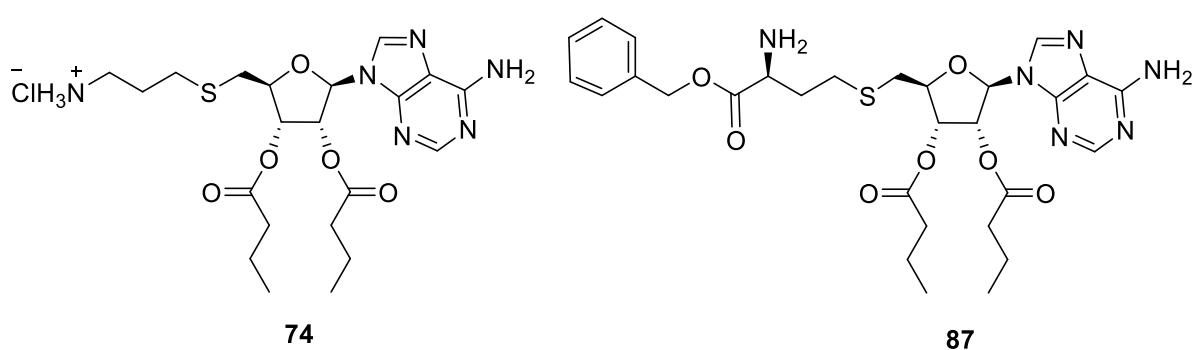


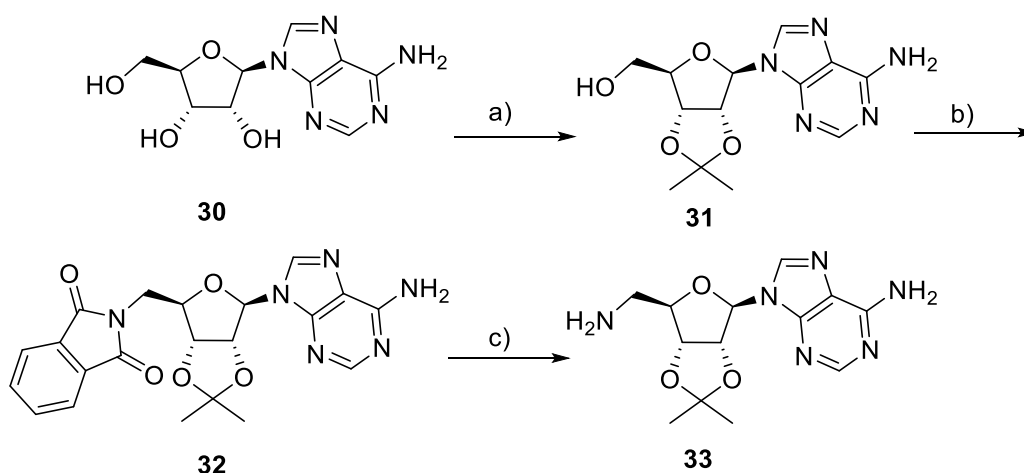
Figure 2-2: Chemical structure of SAM analogues.

*SAM analogues with the hydrophobic version: containing *n*-butyl esters at the 2' and 3' hydroxyls of adenosine 74, 87.*

2.1.1. Inhibitor synthesis

Several AzaSAM analogues have been proposed and designed which combined the structural features of SAM and the substrate arginine motif. Decent inhibition activity was observed for previously synthesised PRMT inhibitors, and that inspired the development of novel bisubstrate inhibitors designed to have superior properties. We were particularly interested in replacing the guanidinium group since this highly basic functional group is protonated by physiological pH and thus carries a permanent charge.

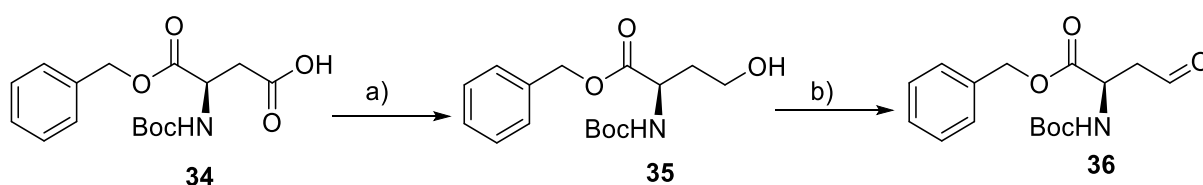
Methods previously developed by the Dowden group were utilised as a starting point for these studies. The readily available protected 5'-amino-5'-deoxyadenosine derivative **30** was generated by the Gabriel synthesis originally described by Townsend and co-workers (Scheme 5).⁵⁴ Adenosine **30** was protected as acetonide **31** with a 85 % yield. Compound **31** was then subjected to Mitsunobu conditions to convert the protected adenosine **31** into phthalimide **32** in good yield (89 %). Intermediate **32** was then treated with hydrazine hydrate in ethanol to give the protected 5'-amino-5'-deoxyadenosine **33** with a 68 % yield.⁵⁴



Scheme 5: Synthesis of 5'-amino-5'-deoxyadenosine **33**. Reagents and conditions: a) TsOH, 2,2-dimethoxypropane, acetone, 85 %; b) DIAD, phthalimide, Ph_3P , THF, r.t., 89%; c) hydrazine hydrate, EtOH, reflux, 68 %.

Attention next turned to synthesise aldehyde **36** which was required for the first reductive amination with starting material **33**. Previously, the amino acid part of the inhibitor was generated from an *N*-Boc protected aspartic acid featuring a *tert*-butyl ester, but this required several synthesis steps.³ An alternative route was to generate aldehyde **36** as a coupling partner for inhibitor synthesis from

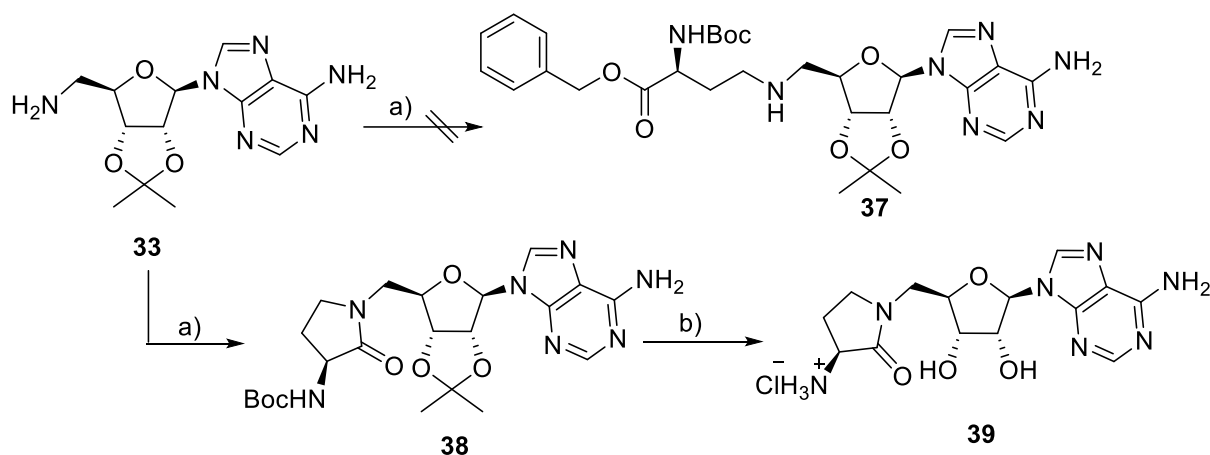
commercially available aspartic acid derivative **34**, which featured a benzyl ester (Scheme 6). Protected aspartic acid **34** was initially treated with *N*-methylmorpholine and isobutyl chloroformate to produce the corresponding mixed anhydride which was reduced *in situ* with sodium borohydride to give an intermediate alcohol **35** with a 90 % yield.⁶⁶ Oxidation of the resulting alcohol **35** was achieved using Dess–Martin periodinane to provide the aldehyde **36** with a 86 % yield (Scheme 6).⁶⁷



Scheme 6: Preparation of aldehyde **36**. Reagents and conditions: a) *N*-methylmorpholine, isobutyl chloroformate, NaBH_4 , 90 %; b) Dess–Martin periodinane, 86 %.

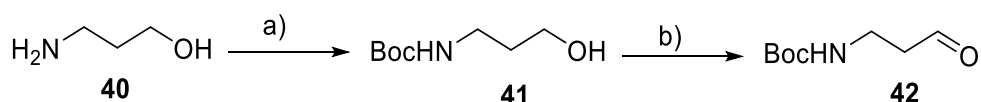
We attempted to use our reductive amination protocol by combining 5'-amino-5'-deoxyadenosine **33** with aldehyde **36** using sodium triacetoxyborohydride in 1,2-dichloroethane. Unfortunately, the desired product **37** was not obtained; instead, intramolecular acylation gave lactam **38** with a 57 % yield. Although this was not what we wanted, we decided to deprotect the acetonide and Boc group in order to generate lactam **39**; thus lactam **38** was dissolved in 90 % (v/v) neat TFA in water to give compound **39**, which was purified using ion-exchange chromatography on Amberlite IRA 400 (Cl^-) to provide the product as pale-yellow oil hydrochloride salt with a 71 % yield (Scheme 7).

Instead, we decided to change the order of the reductive amination steps. In this way, a cheaper aldehyde would be incorporated first, before the amino acid was reacted, which are described below (Scheme 8).



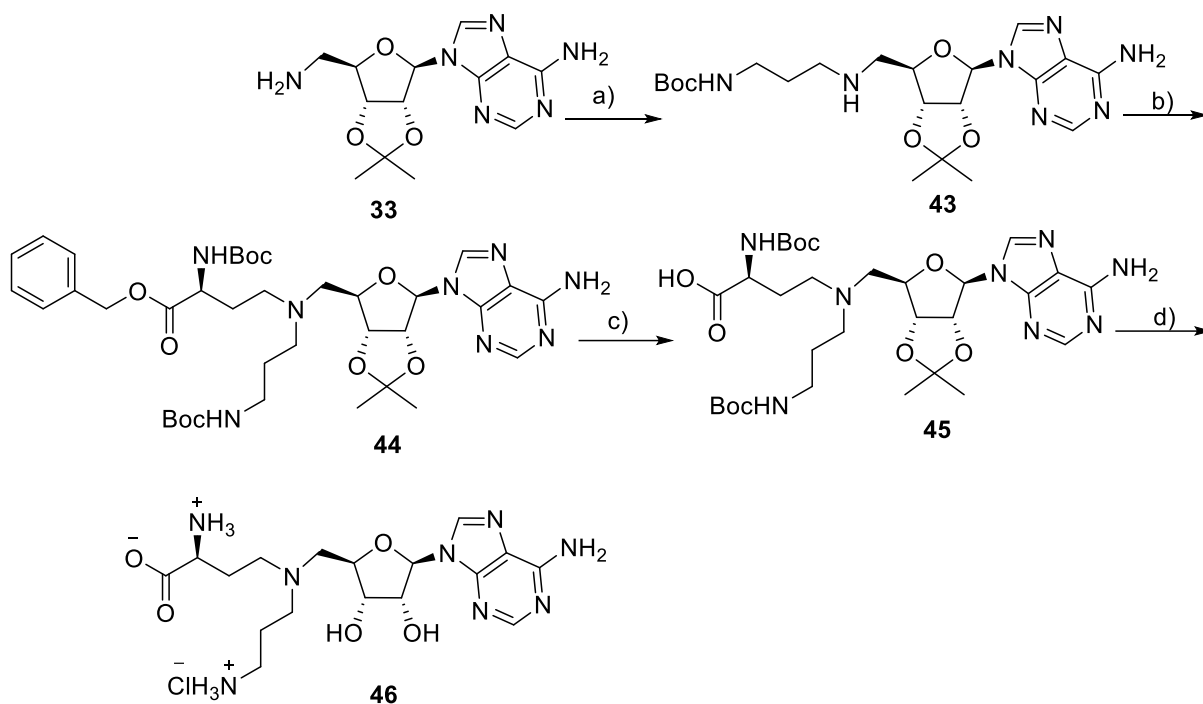
Scheme 7: Synthesis of lactam **39**. Reagents and conditions: a) aldehyde **36**, $\text{NaBH}(\text{OAc})_3$, $\text{ClCH}_2\text{CH}_2\text{Cl}$, r.t., 57 %; b) TFA, H_2O , Dowex 1 $\times$ 2-100 (OH^-) or Amberlite IRA 400 (Cl^-) ion 71 % yield.

Aldehyde **42** was prepared according to a literature procedure.^{68,69} The Boc-protected propylamino aldehyde **42** was used for reductive amination and was made in 2 steps from 3-aminopropan-1-ol. Briefly, a solution of 3-aminopropanol **40** was reacted with di-*tert*-butyl dicarbonate in dichloromethane at room temperature to yield the desired Boc protected alcohol **41** with a 78 % yield.⁶⁸ Oxidation of the remaining alcohol was achieved using Swern oxidation conditions to give the aldehyde **42** with a 85 % yield (Scheme 8).⁶⁹



Scheme 8: Preparation of Boc-protected amino-aldehyde **42**. Reagents and conditions : a) (Boc)₂O, Et₃N, CH₂Cl₂, reflux, 78 %; b) DMSO, oxalyl chloride, Et₃N, CH₂Cl₂, -78 °C, 85%.

With aldehyde building blocks in hand, AzaSAH analogue **43** was prepared; thus 5'-amino-5'-deoxy adenosine **33** was combined with aldehyde **42** in the presence of sodium triacetoxyborohydride as a reducing agent at room temperature to give the corresponding AzaSAH analogue **43** with a 66 % yield.⁵⁹ Interestingly, reductive alkylation between secondary amine **43** and aldehyde **36** using sodium triacetoxyborohydride as the reducing agent gave the desired analogue **44**, without any lactamisation. Presumably, intramolecular acylation would generate an unstable product. Deprotection of desired analogue **46** was performed in two steps. Initially, compound **44** was treated with sodium hydroxide in methanol to give the carboxylic acid **45** with a 68 % yield. Subsequently, this product was dissolved in 90 % (v/v) neat TFA and the resulting concentrated reaction mixture purified using an ion-exchange column to provide the novel inhibitor hydrochloride salt **46** with a 80 % yield (Scheme 9).



Scheme 9: Synthesis of bisubstrate inhibitor **46**. Reagents and conditions: a) Aldehyde **42**, NaBH(OAc)₃, ClCH₂CH₂Cl, r.t., 66 %; b) aldehyde **36**, NaBH(OAc)₃, ClCH₂CH₂Cl, r.t., 75 %; c) 2M NaOH, MeOH, 68 %; d) TFA, H₂O, Dowex 1× 2-100 (OH⁻) or Amberlite IRA 400 (Cl⁻) ion 80 % yield.

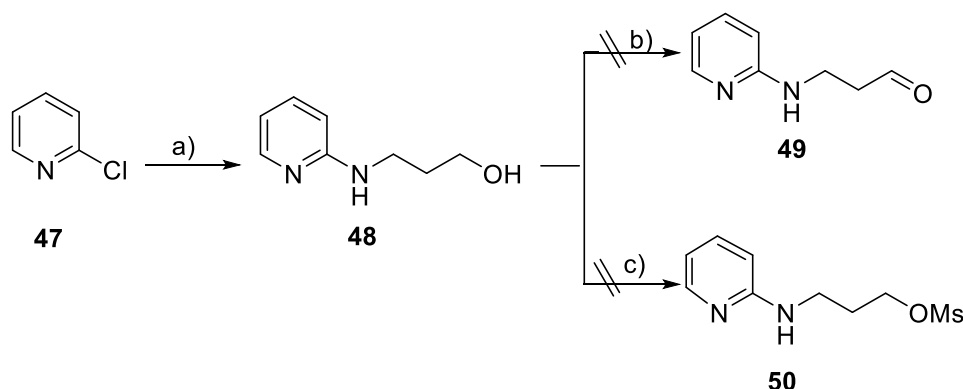
2.1.2. SAM analogues containing hydrophobic isostere 2-aminopyridine

Research then focused on the preparation of secondary amines which possess a guanidinium isostere, 2-aminopyridine **53** (Scheme 12), for incorporation into planned inhibitors. This change in functionality may eventually lead to inhibitors which possess an enhanced cell permeability or may engage in hydrophobic interactions with the residues that line the substrates arginine binding cavity.

Initially, the route was started by the syntheses of a pyridine bearing alcohol **48**, generated in good yields by heating 2-chloropyridine **47**, with a suitable amino alcohol.⁷⁰ Unfortunately, attempts to obtain aldehyde **49** using either Swern or

Dess-Martin periodinane oxidation methods were not successful.^{69,71} Also attempted was the production of a mesylate product **50** through the treatment of the alcohol with methanesulfonyl chloride in the presence of triethylamine in dichloromethane, but unfortunately, only starting materials were returned (Scheme 10), so no attempt to optimise this reaction was undertaken.

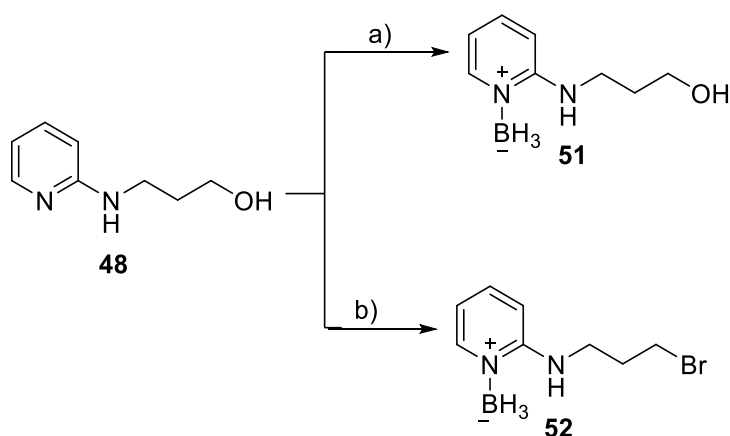
The failed synthesis of compounds **49** and **50** lead us to attempt a modified procedure involving treatment of protected pyridine nitrogen by complexation with borane to reach the target compound, as described below.



Scheme 10: Attempted preparation of aldehyde **49** and mesylate product **50**. Reagents and conditions: a) 2-chloropyridine, 3-aminopropan-1-ol, 150 °C, 67%; b) Dess-Martin periodinane or DMSO, oxalyl chloride, CH₂Cl₂; c) MeSO₂Cl, Et₃N, CH₂Cl₂.

One problem with this strategy is that pyridine is nucleophilic and may undergo a competitive reaction. As an alternative, it was decided to introduce a group which would give temporary protection and one way to achieve this would be to form a stable complex between a pyridine and a Lewis acid, such as BF₃·OEt₂. Evaluation of the literature revealed a patent that suggested that coordination to borane would be effective.⁷² Pyridine-bearing alcohol **48** was reacted with borane

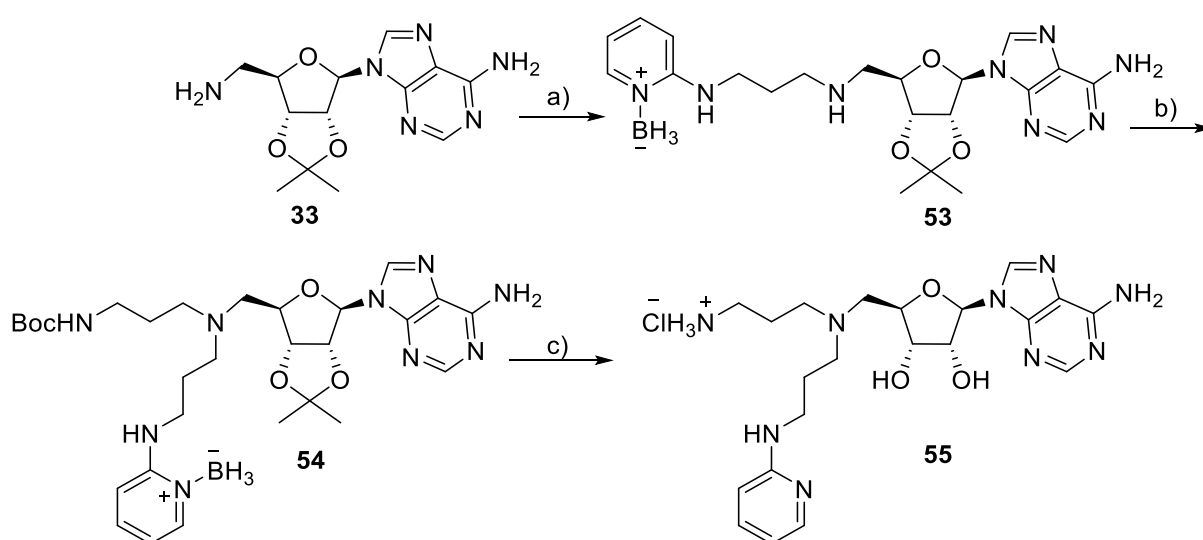
dimethyl sulfide complex in THF (2M) to produce **51** with a 72 % yield, which was stable to chromatography on silica gel. Alternatively, molecule **48** was generated and reacted with thionyl bromide, and, without isolation, the reaction mixture was then treated with *N,N'*-diisopropylethylamine and borane dimethylsulfide complex to provide compound **52** with a 76 % yield (Scheme 11).⁷²



Scheme 11: Preparation of borane-pyridinium bearing alcohol **51** and bromide **52**. Reagents and conditions: a) $(\text{CH}_3)_2\text{S} \cdot \text{BH}_3$ in 2M THF, 72 %. b) SOBr_2 , $(\text{CH}_3)_2\text{S} \cdot \text{BH}_3$, *N,N'*-diisopropylethylamine, CH_2Cl_2 , -60°C , 76 %.

Subsequently, the desired AzaSAM analogue **53** was obtained by alkylation of 5'-amino-5'-deoxy adenosine **33** with bromide **52**, in the presence of potassium carbonate and tetrabutylammonium iodide as a phase-transfer catalyst with a 52 % yield (Scheme 12). When this ion-exchange, the desired mass ion of borane-pyridinium complex of compound **53** was not observed by mass spectrometry. Interestingly, evidence from ^1H and ^{11}B NMR spectrometry indicated that borane was still present in the compound. However, it is possible that borane could be affected by impurities when competing for ionisation or that it was unstable to mass spectrometry as a crude preparation. Subsequently, reductive amination of

secondary amine **53** with aldehyde **42** proceeded using standard conditions to provide the tertiary amines **54** with a 61 % yield. This reaction proceeded quickly, and the desired tertiary amines **54** with stable borane-pyridinium complex was observed as identified by ^1H and ^{11}B NMR and mass spectrometry of the crude preparation. Finally, the inhibitor compound **55** was obtained by deprotection of product **54** by using 90 % (v/v) TFA and purified with Amberlite IRA 400 (Cl^-) ion exchange resin to produce the products as hydrochloride salts, at a good yield (Scheme 12).

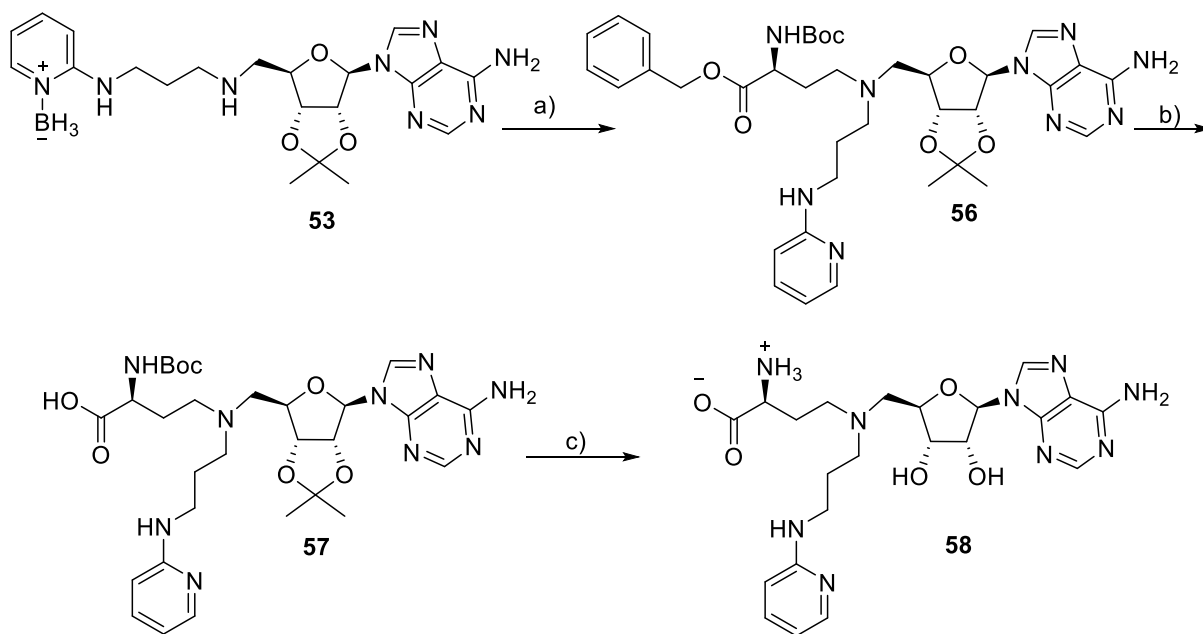


Scheme 12: Synthesis of inhibitor containing hydrophobic isostere 2-aminopyridine **55**.
Reagents and conditions: a) **52**, K_2CO_3 , $(\text{C}_4\text{H}_9)_4\text{NI}$, CH_3CN 52 %; b) aldehyde **42**, $\text{NaBH}(\text{OAc})_3$, $\text{ClCH}_2\text{CH}_2\text{Cl}$, 61 %; c) TFA, H_2O , Dowex 1 $\times$ 2-100 (OH^-) or Amberlite IRA 400 (Cl^-) ion 80 % yield.

AzaSAM analogues **58** (Scheme 13) were also designed, which contained a terminal amino acid methionine instead of propylamine group, to represent the most hydrophobic analogues amongst the proposed targets. Thus, it was decided

to synthesise AzaSAM analogues **58** to investigate whether our target enzyme could exploit similar reactivity.

With the key intermediate **53** in hand, AzaSAM analogues **58** containing an aminopyridine group was approached. Thus, reductive amination of the secondary amine **53** with the aspartic acid derived aldehyde **36** using sodium triacetoxyborohydride in 1,2-dichloroethane provided the amino acid derived precursor **56** with a 56 % yield. The borane-pyridinium complex of compound **56** proved unstable and mostly decomposed before full characterisation could be achieved. This assignment was confirmed using ^1H and ^{11}B NMR spectrometry as well as mass spectrometry. Subsequently, the benzyl group was removed by hydrolysis using 2M aqueous sodium hydroxide to give compound **57** at a good yield. Deprotection of the compound **57** was carried out by using 90 % (v/v) TFA in water, followed by ion-exchange chromatography on Amberlite IRA 400 (Cl^-) to give the novel inhibitor **58**, as a pale-yellow hydrochloride salt (Scheme 13).



Scheme 13: Synthesis of inhibitors which contain hydrophobic isostere 2-aminopyridine **58**. Reagents and conditions: a) aldehyde **36**, $\text{NaBH}(\text{OAc})_3$, $\text{ClCH}_2\text{CH}_2\text{Cl}$, 56 %; b) 2M MeOH, 68 %; c) TFA, H_2O ; the Dowex 1 $\times$ 2-100 (OH^-) or Amberlite IRA 400 (Cl^-) ion 83% yield.

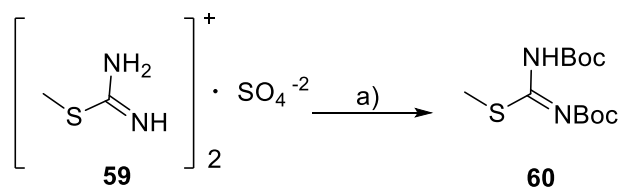
Compounds **55**, **58**, and **46** were taken forward for evaluation in a methylation assay and evaluation of binding with CARM1 or PRMT1, using isothermal titration calorimetry (ITC) and *in vitro* methylation assay as discussed in more detail later in chapter three (see Section 3.2.4. and 3.3.6.).

Additionally, co-crystallisation of inhibitors **55** and **58** with CARM1 revealed that these compounds occupied the arginine and SAM binding cavities (see Section 4.3.6.3). These two compounds showed an *in vitro* methylation assay selectivity toward CARM1 over PRMT1, and these results encouraged the synthesis of similar bisubstrate inhibitors containing hydrophobic group such as pyridine for further future work.

2.1.3. Nitrogen SAM analogue containing guanidine.

It was decided to carry out the synthesis of a bisubstrate inhibitor containing guanidinium group **64** (Scheme 15),⁶² which had previously been reported by another group. This approach allowed for a comparison between its bindings to our bisubstrate inhibitors with other protein arginine methyltransferases being prepared during the project.

Initially, boc-protection of S-methylisothiourea sulfate **59** gave guanidinylation agent **60** in quantitative yield from in a mixture of saturated aqueous Na₂CO₃ and dichloromethane (Scheme 14).⁷³

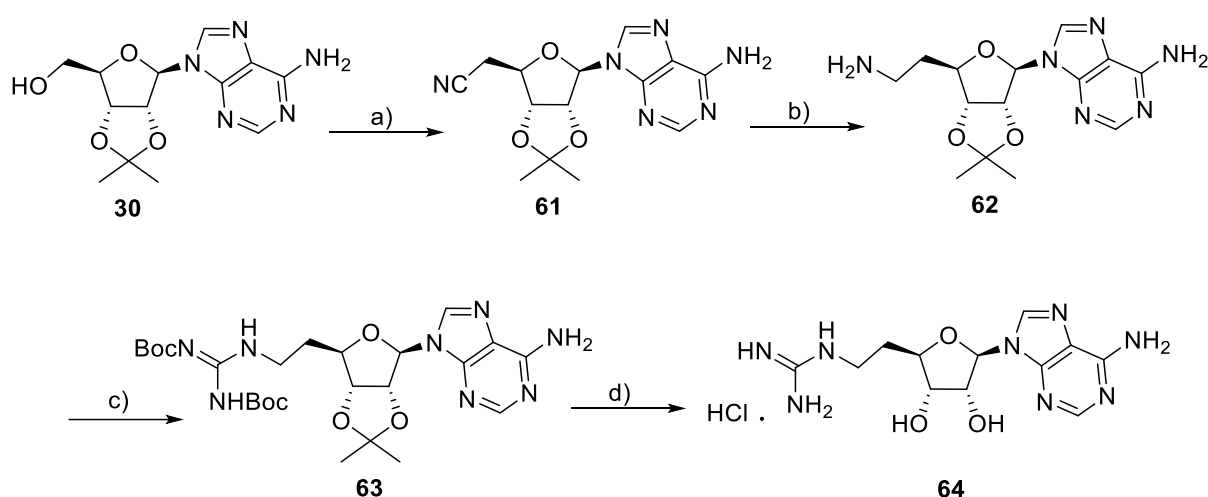


Scheme 14: Preparation of guanidinylation agent **60**.

Reagents and conditions: a) (Boc)₂O, sat. Na₂CO₃(aq)/CH₂Cl₂, r.t., quant.

The target inhibitor **64** was synthesised from starting material **30** using previously reported procedures with an overall yield of 65 % produced using four steps (Scheme 15).⁶² Briefly, a Mitsunobu reaction was used to convert the protected adenosine **30** into nitrile **61**; thus, treatment of adenosine **30** with PPh₃, DEAD, and acetone cyanohydrin gave nitrile **61** with a 60 % overall yield. Subsequent reduction of intermediate nitrile using platinum dioxide (PtO₂), AcOH and H₂ successfully converted **61** into amine **62**.⁶² As soon as these steps had been carried out, intermediate amine **62** was then heated with protected

methylisothiurea **60** at 50 °C for 2 hours to produce the corresponding protected guanidine species **63** at a satisfactory yield.⁶² Subsequent global deprotection of protected guanidine compound with 2M HCl in diethyl ether gave the respective prototype inhibitor **64** as the pale-yellow hydrochloride salt in approximately 65% to quantitative yield, after ion-exchange chromatography (Scheme 15).



Scheme 15: Synthesis of inhibitor containing guanidine **64**. Reagents and conditions: a) **30**, acetone cyanohydrin, PPh_3 , DEAD, THF, 90 %; b) PtO_2 , AcOH, H_2 , quant.; c) N,N' -di-Boc- N'' -triflyl-guanidine, NEt_3 , CH_2Cl_2 , 75 %; d) 2M HCl in diethyl ether 65 %.

Overall, AzaSAM analogues were successfully synthesised by a route involving Gabriel synthesis of protected 5'-amino-5'-deoxyadenosine **33** and successive reductive aminations with suitable aldehydes. The challenge of generating the 2-aminopyridine derivatives was met by employing a stable borane-pyridinium complex, as an unusual protecting group. This strategy allowed the same modular reductive amination sequence to be applied to the synthesis of the novel bisubstrate inhibitors **55** and **58**.

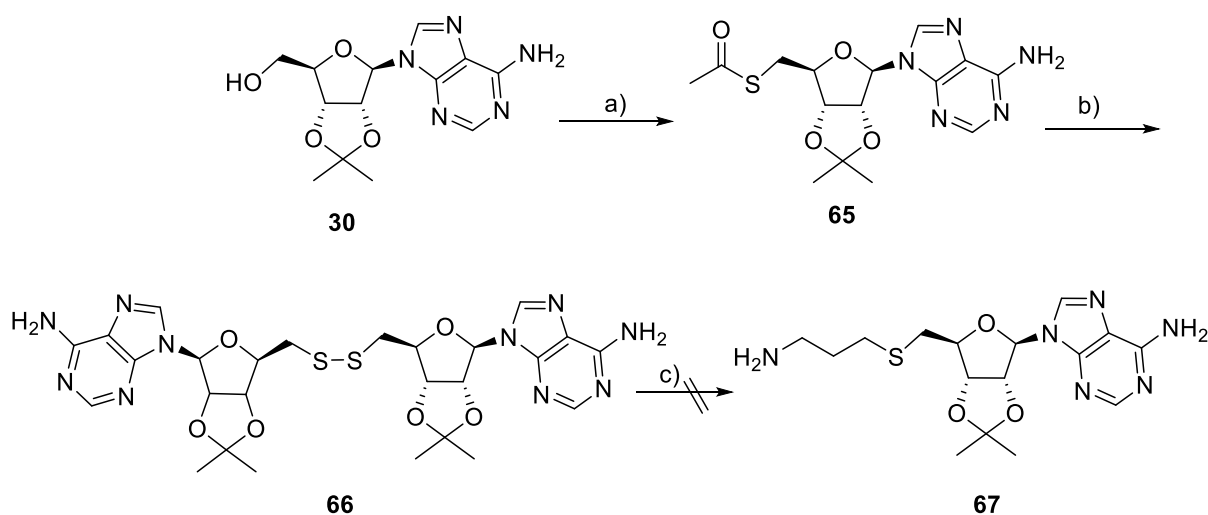
2.2. SAM analogue containing butyl ester on 2' and 3' of adenosine

As described in chapter 1, *S*-Adenosyl methionine is the second most abundant co-factor in the human body after adenosine-5'-triphosphate (ATP). There is much interest in developing SAM analogues that bear reactive functionality other than a methyl group at the sulfonium centre since these would transfer a distinct label to proteins. Meanwhile, bump-and-hole strategies have been developed in which structurally modified co-factor analogues are developed to be recognised by the complementary binding site of a mutant protein, in this way selective labelling of specific methyltransferases inhibitors might be achieved.^{74,75,76} The problem with these approaches is that SAM is not cell-permeable,^{65,64} so complementary biosynthetic machinery is required to accommodate production of these modified co-factors within cells. A simpler approach would be to make SAM cell permeable; thus we set out to make SAM compounds **74** and **87** with additional *n*-butyl esters at the 2',3' hydroxyl of SAM with a view to examining whether they are able to cross the membrane.^{65,77} In principle, the butyl ester would then be hydrolysed to deliver SAM within the cell.

2.2.1. Synthesis of the model compound

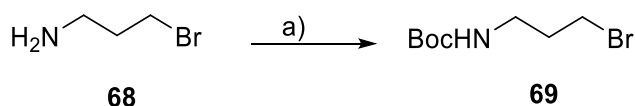
The strategy for synthesising the target compound was to use 5'-deoxy-5'-thioadenosine **66** as a building block. The approach was based in part on the success of inhibitors containing the adenosine fragment rationally generated from the thioadenosine **66** intermediate.⁷⁸ Therefore the developed methodology was used to synthesise a range of related SAH analogues containing butyl esters at the 2',3' hydroxyl group which could, in turn, possess exciting biological activity.

A suspension of 2',3'-O, O-isopropylideneadenosine **30** was synthesised in a single step from adenosine as described in Section 2.1.1., Scheme 5. This compound was used as a starting point to generate the protected thionucleoside **66**. Protected adenosine derivative **30** was subjected to Mitsunobu reaction by treating with thioacetic acid and triphenylphosphine in THF to afford thioacetate compound **65** with an excellent 93 % yield.^{78,79} Subsequent removal of the acetyl group by addition of methanol saturated with ammonia gave disulfide **66** with an excellent 95 % yield. This compound gave major peaks at m/z 324, and at m/z 645 in the mass spectrum of the reaction mixture; the peak at m/z 645 was a result of the disulfide formation, which has been reported previously (Scheme 16).⁷⁹ *In situ* reduction of disulfide **66** with sodium borohydride and subsequent alkylation with commercially available 3-bromopropylamine hydrobromide, in the presence of potassium carbonate, in DMF was attempted. Although the target compound **67** was observed using mass spectrometry analysis of the crude reaction mixture, attempts to extract this product from the aqueous layer by using various organic solvents were unsuccessful (Scheme 16).



Scheme 16: Attempts to synthesise compound **67**. Reagents and conditions: a) **30**, DIAD, PPh_3 , thioacetic acid, THF, 0 °C 93 %; b) ammonia in methanol, r.t., 95%; c) K_2CO_3 , $NaBH_4$, DMF, 100 °C.

In the light of the unsuccessful attempt to produce the compound **67**, we sought to protect the nitrogen of 3-bromopropylamine **68**. 3-Bromopropylamine **68** was treated with di-*tert*-butyl dicarbonate in the presence of Et_3N in CH_2Cl_2 to afford the crude protected **69** with an excellent yield of 92 % (Scheme 17).

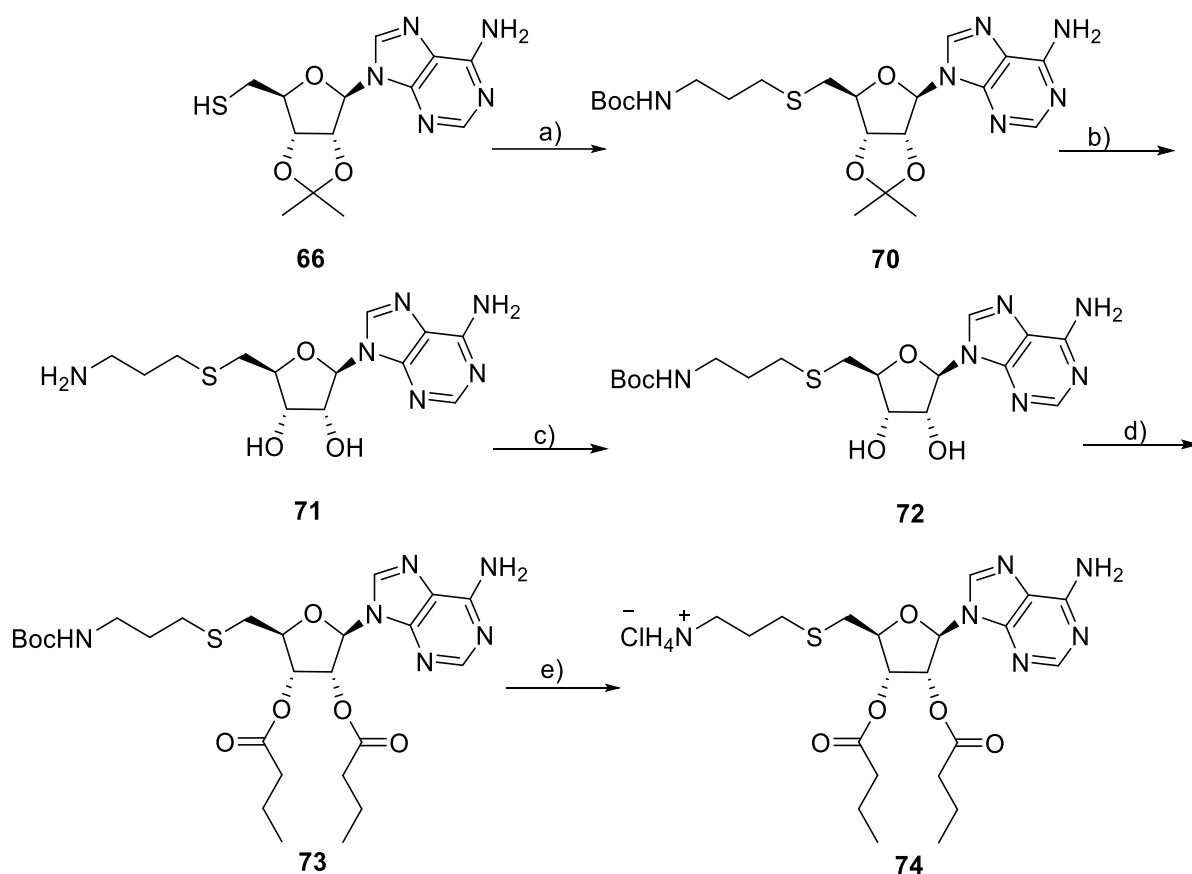


Scheme 17: Preparation of Boc-protected amino-bromide **69**. Reagents and conditions: a) $(Boc)_2O$, Et_3N , CH_2Cl_2 , r.t., 92 %.

Next thiol **66** was alkylated with bromide **69** in the presence of sodium borohydride and potassium carbonate, followed by heating at 100 °C to give analogue **70** (Scheme 18) with a 63 % yield (sodium borohydride was added to

the mixed reaction to ensure that the disulfide was reduced during the reaction). Secondary alcohols needed to be released before subsequent acylation but unfortunately, attempts to selectively remove the acetonide in the presence of the Boc group, to produce compound **71**, were not successful (Scheme 18). Instead, both the acetonide and the Boc group were cleaved using 90 % (v/v) TFA in water to provide compound **71**, but some impurities, which included other adenine-containing compounds contaminated the residue. The residue was applied to Amberlite IRA 400 (Cl⁻) ion exchange resin, eluting with water and methanol, but unfortunately, there was still some impurities contaminating the product. Attempts to use other resins like Dowex 1× 2-100 (OH⁻) ion-exchange resin also failed to remove purities. Purification of the highly polar target compound **71** was a significant problem. Thus, the protected compound **72** was produced by treatment of intermediate **71** with Boc anhydride and triethylamine in anhydrous methanol with a 65 % yield (Scheme 18).

Preparing a hydrophobic version of the sulfur-linked compound by adding butyl esters at the 2',3' hydroxyl group of adenosine was explored. The intermediate containing butyl ester on 2' and 3' of adenosine **73** was synthesised by treating intermediate **72** with a butyric anhydride in pyridine to give the diester intermediate. The crude product was treated with 90 % (v/v) neat TFA, in order to remove the Boc protecting group, which produced the free amine **74** with a 83% yield after work up (Scheme 18).



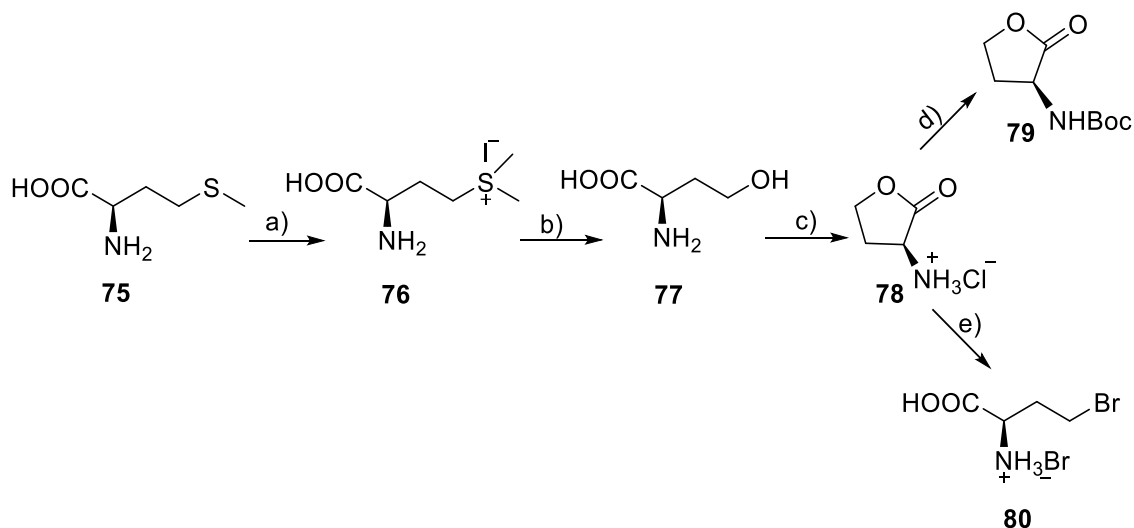
Scheme 18: Synthesis of a SAM analogue which contains a butyl ester on 2' and 3' of adenosine **74**. Reagents and conditions: a) **69** bromide, K_2CO_3 , DMF, $NaBH_4$, 63 %; b) TFA, H_2O , Dowex 1 $\times$ 2-100 (OH^-) or Amberlite IRA 400 (Cl^-) ion; c) **71**, Boc_2O , MeOH, Et_3N , 65%; d) butyric anhydride, pyridine, r.t., 55 %; e) **73**, TFA, H_2O , r.t., 83 %.

Next, it was decided to synthesise SAH analogue **81** (Scheme 20) which contained a terminal α -amino acid moiety instead of propylamine. The primary purpose of this target molecule was to allow the co-factor to be able to pass through the cell membrane. The synthetic route started by generating lactone **79** or bromide **80** as a starting point following a described procedure.⁸⁰

Lactone **79** or bromide **80** were synthesised with a view to generating the SAM analogue **81** (Scheme 20) by alkylating with 5'-thio-5'-deoxy-adenosine **66**. As

illustrated in Scheme 19, L-methionine **75** was treated with iodomethane in water to give sulfonium salt **76** with an excellent yield of 84 %.^{80,81} Then, homoserine **77** was synthesised from this sulfonium salt **76** by reaction with hydroxide using sodium hydrogen carbonate in water with a 53 % yield. Subsequent cyclisation in the presence of 2.4 M aqueous HCl under reflux for 3 hours gave the lactone **78** with a 57 % yield.^{80,81} The crude product **79** was synthesised from lactone **78** with Boc anhydride and triethylamine in a 69 % yield. The idea here was that protecting the lactone for subsequent alkylation with thiol **66** would give a hydrophobic product amenable to work up in the organic solvent.

Alternatively, opening the ring of lactone **78** using hydrobromic acid in acetic acid was attempted, and the desired mass ion was detected in the reaction mixture by mass spectrometry. Unfortunately, scale-up assays gave a complex mixture, and the product was difficult to purify because the bromide **80** was contaminated with the lactone starting material (Scheme 19).

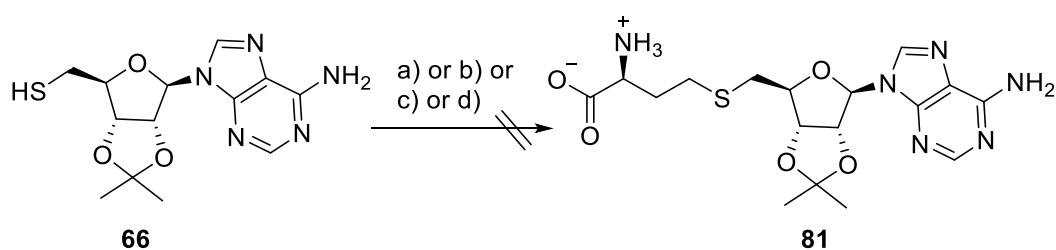


Scheme 19: Preparation of lactone **79** and bromide **80**.

Reagents and conditions: a) MeI, THF, 84 %; b) NaHCO₃, H₂O, reflux, 53 %; c) 2.4 M HCl, H₂O 57 %; d) Boc₂O, Et₃N, DMF, 69 %; e) HBr/CH₃COOH.

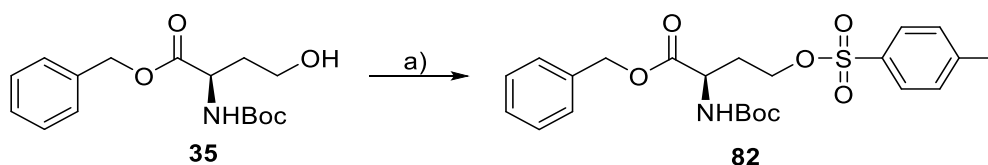
It was hoped that it might be possible to open the lactone ring *via* S_N2 nucleophilic displacement with the thiol group of the compound. Synthesis of SAH analogue **81** was attempted by alkylation of 5'-thio-5'-deoxy-2',3'-O-isopropylidene adenosine **66** with lactone **79** under various conditions (Scheme 20), but unfortunately, no desired product **81** was obtained.

Similarly, disappointing results were obtained from attempted reaction 5'-thio-5'-deoxy-2',3'-O-isopropylidene adenosine **66** and bromide **80** in the presence of K₂CO₃ in DMF, but the zwitterionic crude product remained in the aqueous layer and could not be purified (Scheme 20).



Scheme 20: The unsuccessful SAH analogue **81** reactions. Reagents and conditions: a) **79**, DMSO, NaH, b) **79**, DMF, K₂CO₃, c) **79**, CH₂Cl₂, Et₃N, d) **79**, dioxane, K₂CO₃.

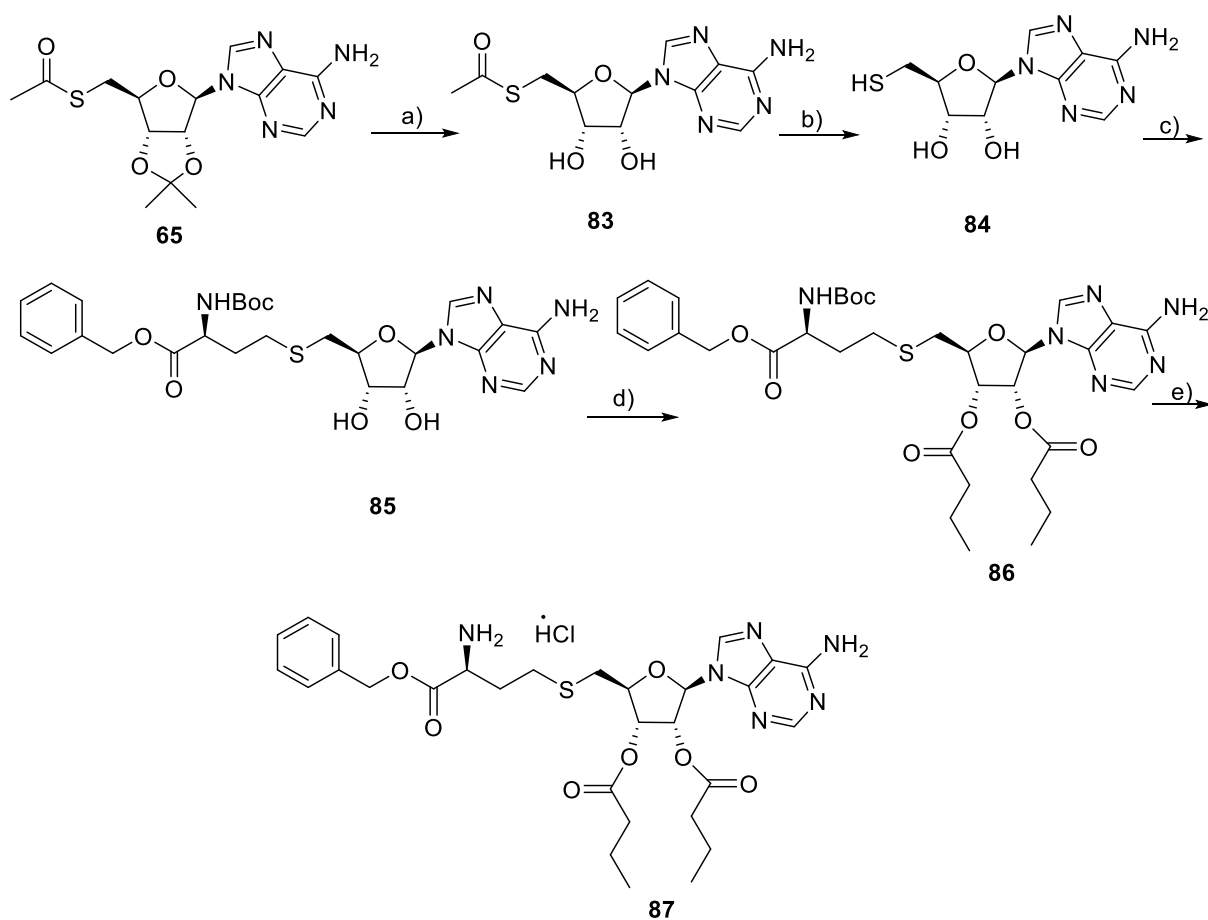
Unfortunately, the target compound **81** discussed above was unstable during purification. Alternative alkylating agents for reaction with 5'-thio-5'-deoxy-2',3'-O-isopropylidene adenosine **66** were sought. Already obtained was alcohol **35** which featured a benzyl ester which would give a hydrophobic product amenable for future work with the organic solvent; thus, treatment of alcohol **35** with p-toluenesulfonyl chloride and pyridine in dichloromethane gave compound **82** with a yield of 80 % (Scheme 21).



Scheme 21: Synthesis of compound **82**. Reagents and conditions: a) *p*-toluenesulfonyl, pyridine, CH_2Cl_2 , 0°C 80 %.

Previously, deprotection of the Boc group during acetonide deprotection (Scheme 18) had been observed, so acetonide **85** was initially removed using aqueous TFA to give 5'-acetylthio-adenosine **83** with an excellent yield of 90 %.⁸² A suspension of 5'-acetylthio-adenosine **83** was then hydrolysed in the same manner as before for compound **66** (Scheme 18) using methanol saturated with ammonia to afford thiol **84** with a 90 % yield. A suspension of **84** was alkylated with tosylate **82** in the presence of potassium carbonate and sodium borohydride (to reduce disulfide) in DMF to give compound **85** with a 60 % yield. The intermediate containing butyl ester was synthesised by treatment of intermediate **85** with a butyric anhydride in pyridine, which installed the required esters at the 2' and 3' of adenosine to give **86** with a 87 % yield. Removal of the Boc protecting group was with 90 % (v/v) TFA, and the free amine **87** obtained at 83 % yield after work up (Scheme 22).

Deprotection of the benzyl ester of desired analogue **87** was performed by treating with sodium hydroxide in methanol with the aim of producing a carboxylic acid. Unfortunately, the butyl esters at the 2',3' hydroxyl and the benzyl ester groups also decomposed. Had time permitted, future work would have involved assays looking at deprotecting the benzyl ester of analogue **87** such as by hydrogenolysis.



Scheme 22: Synthesis of SAM analogue which contains butyl ester on 2' and 3' of adenosine **87**. Reagents and conditions: a) TFA, H₂O; 90 %; b) NH₃ in MeOH, 90 %; c) **82**, K₂CO₃, NaBH₄, DMF 60 %; d) butyric anhydride, pyridine r.t., 87 %; e) TFA, H₂O, r.t., 83 %.

In summary, synthesis of SAH with butyl esters at the 2',3' hydroxyls of adenosine **74** and **87** were successfully performed *via* a Mitsunobu transformation with thioacetic acid. The first synthesis route was attempted by reacting 5'-thio-5'-deoxy-2', 3'-O-isopropylidene adenosines **66** with homoserine lactone **79** (or **80**) but unfortunately no desired compound was obtained. Alternatively, preparation of tosylate derivative **82**, allowed us to alkylate thio-adenosine **84** to give sulfur-linked analogue **85**. Deprotection of these compounds was achieved by successfully removing the Boc group by using TFA with water and was purified by using Amberlite IRA 400 (Cl⁻) ion exchange resins to adjust the counterion for biophysical characterisation.

Improvement of a hydrophobic version of the sulfur-linked compounds **74** and **87** represents an interesting molecule; future biological assays will be required to understand whether they are able to pass through the cell membrane.

3. Chapter 3 Optimising the production, purification and crystallisation of human PRMT1 and CARM1

3.1. Choice of constructs

This chapter is divided into two sections which discuss the production, and purification of PRMT1 and CARM1 constructs, respectively. These constructs were characterised and evaluated for enzymatic activity with a broader aim of comparing the activity and binding of bisubstrate inhibitors.

The first section of this chapter will examine the structure and function of PRMT1 (residues 22-361) using an expression plasmid that was cloned by another member of the wider research group.⁸³ This PRMT1 construct (residues 22-361) was transformed into LOBSTR BL21 (DE3)-RIL cells for the purpose of subsequent protein production (Figure 3-1). This strain is derived from BL21 (DE3) codon plus, which is another *E. coli* strain commonly used for protein production. The overarching objective was to improve the purity of these target proteins. The LOBSTR BL21 (DE3)-RIL strain carries a modification in its ArnA and SlyD genes, the most common contaminants in *E. coli*, which reduces their affinity to the nickel column.⁸⁴

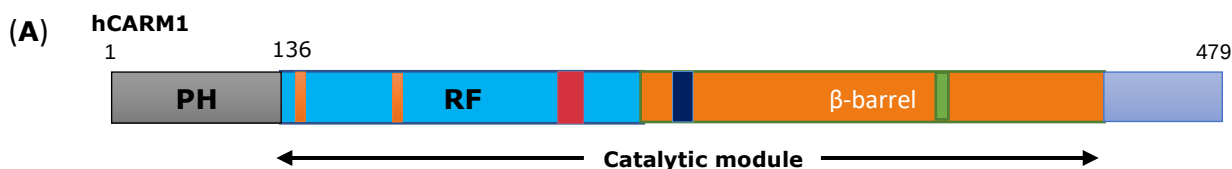
PRMT1 (residues 22-361)

M A A A E A A N C I M E N F V A T L A N G M S L Q P P L E E V S C G Q A E S S E K P N A E D M T S K D Y
Y F D S Y A H F G I H E E M L K D E V R T L T Y R N S M F H N R H L F K D K V V L D V G S G T G I L C M F
A A K A G A R K V I G I E C S S I S D Y A V K I V K A N K L D H V V T I I K G K V E E V E L P V E K V D I I I
S E W M G Y C L F Y E S M L N T V L Y A R D K W L A P D G L I F P D R A T L Y V T A I E D R Q Y K D Y
K I H W W E N V Y G F D M S C I K D V A I K E P L V D V D P K Q L V T N A C L I K E V D I Y T V K V E D
L T F T S P F C L Q V K R N D Y V H A L V A Y F N I E F T R C H K R T G F S T S P E S P Y T H W K Q T V F
Y M E D Y L T V K T G E E I F G T I G M R P N A K N N R D L D F T I D L D F K G Q L C E L S C S T D Y R M
R

Figure 3-1: The sequence of the catalytic domain of PRMT1.

Catalytic domain residues of PRMT1 (residues 22-361) in which residues cloned into the expression vector are highlighted in blue and omitted amino acids are highlighted in red.

The next section of this chapter explores the importance of the stable catalytic domain construct, namely CARM1 (coactivator-associated arginine methyltransferase 1). The full-length CARM1 protein consists of 608 amino acids which are present in two main domains: an N-terminus PH domain connected to the catalytic domain of 350 residues (Figure 3-2). Most reported studies have examined the catalytic domain of CARM1 for studying inhibitor binding and for crystallography. Residues 135-479 of CARM1 protein, representing the catalytic domain, were chosen for these studies based on a published crystal structure (PDB code 2Y1X).⁴¹ It has been reported that full-length CARM1 is challenging to express as a soluble protein in bacteria, but it can be obtained using mammalian or insect expression systems.⁴⁰ Obtaining a full-length CARM1 crystal structure would allow for a greater understanding of the influence of the C-terminal sequence on the active site; additionally, it would inform whether the catalytic domain can act as a suitable surrogate for the full-length protein when studying inhibitor binding to the protein.



(B) CARM1 (residues 136-479)

MAAAAAVGPAGGAGSAVPGGAGPCATVSVFPGARLLTIGDANGEIQRHAE
 QQALRLEVRAGPD SAGIALYSHEDVCV FKCSVSRETECSRVGKQSFII TLGCN
 SVLIQFATPND FCSFY NIKTKRGHTLERSVFSERTEESS AVQYFQFYGY LSQ
 QQNM MQDYV RTGTYQ RAILQNHTDFKDKIVLDVG CG SGILS FFAAQAGARKI
 YAVEASTMAQHA EVLVKSNNTDR IVVIPGKV EE VSLPEQVD I ISEPMGYML
 FNERMLESYLHAKKYLKPSGNMFPTIGDVHLAPFTDEQLYMEQFTKANFWYQ
 PSFHGVDLSALRGA AVDEYFRQPVV DTFDIRILMAKSVKYTVN FLEAKE GD L
 HRIEIPFKF HMLHSGLVHGLAFW FDVAFIGSIMTVWLSTAPTEPLTHWYQ VRC
 LFQSPLFAKAGDTLSGTCLL IANKR QSYDISIVA QVDQTG SKSSNLLDLKN PF
 FRYTGTT P S P P P G S H Y T S P S E N M W N T G S T Y N L S S G M A V A G M P T A Y D L S S V I A
 S G S S V G H N N L I P L A N T G I V N H T H S R M G S I M S T G I V Q G S S G A Q G S G G G S T S A H
 Y A V N S Q F T M G G P A I S M A S P M S I P T N T M H Y G S

Figure 3-2: Overall domain architecture of human CARM1 (residues 135-479).

A) The top panel represents the architecture of the PH and catalytic domain residues of human CARM1. B) CARM1 sequence 135-479, where catalytic core domain residues are highlighted in blue, while omitted amino acids are shown in red

3.2. PRMT1: purification and crystallisation

3.2.1. Choice of construct

Rat PRMT1 (residues 14-353) has been successfully crystallised and shown to be catalytically active *in vitro*.²⁴ Rat PRMT1 has been shown to have comparable *in vitro* methyltransferases activity to human PRMT1.²⁶ Human PRMT1 (residues 22-361) had already been cloned within our group by Dr Emma Gunnell.⁸³ This construct was chosen based on the sequence and structural similarity to the rat PRMT1 (residues 14-353) catalytic domain. The aim was to use this protein for binding and inhibition studies and to investigate crystallisation of the protein when conjugated with putative bisubstrate inhibitors.

3.2.2. PRMT1 production and purification

PRMT1 domain (residues 22-361) in pET26b was produced and purified using the six histidine residues present at the N-terminus.²⁶ The protein obtained appeared highly pure after nickel affinity chromatography and eluted from the gel filtration in one main peak (Figure 3-3). The fractions that were subsequently shown to contain PRMT1 (40.4 kDa) were always observed close to the void volume of the gel filtration column at (~55 mL), which indicates a mass of ~440 kDa. This behaviour suggests higher molecular weight aggregates or oligomers. PRMT1 has been shown to form oligomers which exceed 600 kDa,²⁴ proteins of this size are at the upper size limit for proteins resolvable using a HiLoad 16/600 Superdex 200 pg gel filtration column. It was, therefore, difficult to distinguish using this experimental approach between PRMT1 aggregates and high molecular weight oligomers. Analysis of the SDS-PAGE gel under denaturing conditions was consistent with PRMT1 (MW=40.4 kDa), but a small band was observed at around 80.0 kDa, which suggests protein contamination (Figure 3-3). The total yield of PRMT1 was 2.75 mg/L of bacterial culture; this yield was sufficient to allow activity assays with our bisubstrate inhibitors and to facilitate crystallisation trials. In terms of protein activity, the protein was found to be active using an *in vitro* methylation assay, which has been optimised by Dr Emma Gunnell.⁸³

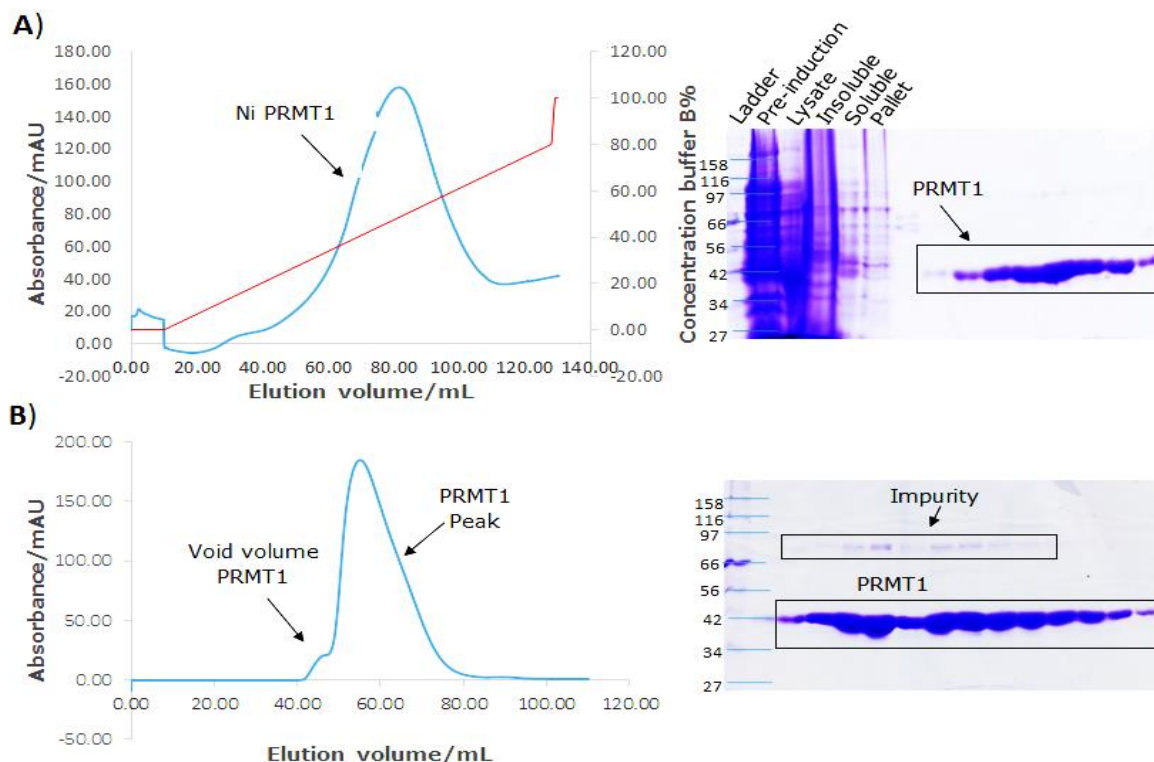


Figure 3-3: Purification of PRMT1 with a His-tag at the N-terminus.

A) Nickel affinity column purification chromatogram showing a UV trace (blue) corresponding to proteins eluted based on the concentration of percentage buffer B (red). The corresponding SDS-PAGE analysis revealed one peak, which eluted between 45-90 mL. The fractions arising from nickel column purification were combined and concentrated for subsequent gel filtration. B) Gel filtration (Superdex 200 16/60) chromatogram from hPRMT1 purification with the corresponding SDS-PAGE. PRMT1 eluted from the gel filtration column at 55 mL as an asymmetric peak with long-tail corresponding to a MW of > 440 kDa. The corresponding SDS-PAGE analysis of the gel filtration fractions shows the presence of a faint band corresponding to an unidentified protein with higher mass ~80.0 kDa.

As a complementary means for determining the potency of the inhibitors, we sought to use an *in vitro* methylation assay to determine whether the selected bisubstrate inhibitors inhibited the target enzyme (PRMT1) (Figure 3-4), expressed as IC₅₀ values. This assay was used for its simplicity and diverse substrate compatibility. In order to do this, it was necessary to produce substrate

peptides that could be recognised and methylated by CARM1 and PRMT1 enzymes. It was previously observed that PRMT1 is able to methylate histone 4 on arginine 3 (H4R3)⁸⁵ while CARM1 methylates histone 3 on arginine 17 and 26 (H3R17 over H3R26), respectively.⁸⁶ Histone 4 peptide (residues 1-22, as GB-fused peptide) for PRMT1 and histone 3 peptide (residues 3-31, as GB-fused peptide) for CARM1 were already sub-cloned into GB-pET21a, within our group.⁸³

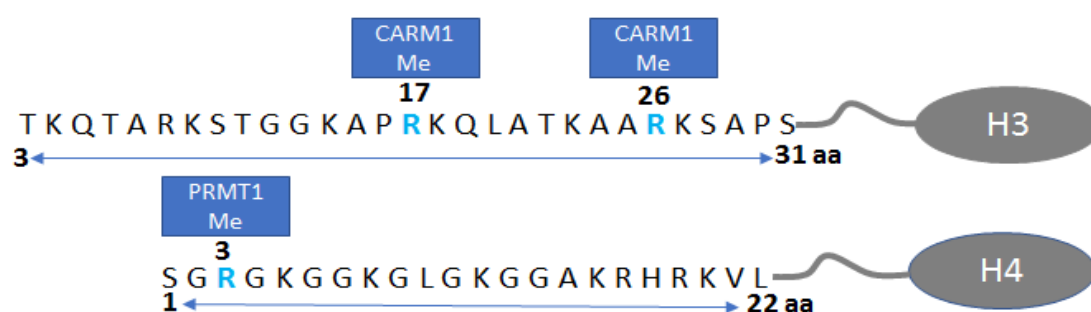


Figure 3-4: Sequence representation of the peptides from histone 3 and 4.

The blue squares represent the enzymes of PRMT1 and CARM1 methylated arginine residues highlighted in light blue which lie within the H3 and H4 boundaries respectively. The aa sequences for each peptide (residues 1-22 H4 and residues 3-31 H3) were cloned into the expression vector. Cloning of this construct was performed by Dr Emma Gunnell.

These peptides were produced and purified as GB-fusions (data not shown) and could then be used as substrates to test the inhibition of compounds against both PRMT1 and CARM1. The 6.0 kDa GB tag is the B1 domain of streptococcal protein G (PDB code 3GB1).⁸⁷ A modified pET21a vector was used that encoded the B1 domain of the streptococcal protein (GB (6.0 kDa) for use as an N-terminus fusion tag that can aid the production of peptides and allow visualisation of the substrate using SDS-PAGE. Optimisation of the assay using PRMT1 and this peptide is detailed in the following section, while optimisation of CARM1 reactions is further described in Section 3.3.

3.2.3. Evaluation of selected compounds as inhibitors of PRMT1 by *in vitro* methylation assay

As mentioned in chapter 2 (see section 2.1), related amine-guanidine inhibitors were found to be inhibitors of PRMT1.^{2,3} We next examined bisubstrate inhibitors binding to PRMT1 using isothermal titration calorimetry (ITC) and inhibition using an *in vitro* methylation assay.

Five potential bisubstrate inhibitors (**55**, **58**, **46**, **64**, and **39**) were initially evaluated using an *in vitro* methylation assay against PRMT1. Briefly, various concentrations of each compound were incubated in the presence of recombinant PRMT1, histone H4 substrate (residues 1-21) and [³H]-SAM for 60 min, at 25 °C. The target protein was then isolated from the reaction mixture by binding it to a phosphocellulose membrane, the membrane was then washed with buffer for 3 × 10 minutes to remove any unbound [³H]-SAM, and the incorporation of the [³H]-methyl group into the remaining protein fraction measured by counting scintillation (these experiments were carried out by Dr Emma Gunnell, see Appendix 9.1).

The enzymatic activity of PRMT1 was determined by using an *in vitro* methylation assay, wherein a substrate peptide derived from the histone H4 tail at 40 µM concentration was treated with [³H]-SAM 0.23 µM as the methyl donor and PRMT1 160 nM. Bisubstrate inhibitors at varying concentrations (1 mM, 100 µM and 10 µM) were used to determine the minimum inhibition concentration of PRMT1. These studies showed that the general methyltransferase inhibitor, aminopropyl derived azaSAM **46**, was the most potent inhibitor because the presence of 10 µM reduced PRMT1 activity by approximately 60 % (Figure 3-5). The 2-aminopyridine

derived compounds **55** and **58** were weaker inhibitors of PRMT1, assays using 10 μ M of either compound reduced PRMT1 activity by approximately 40 % and 10 %, respectively. Similarly, 10 μ M of compound **64**, which has previously been reported as an inhibitor of CARM1,⁶² reduced PRMT1 activity by 10 %. In contrast, lactam **39**, which does not resemble AzaSAM structurally, reduced PRMT1 activity by approximately 5 % in assays using 10 μ M and by 20 % when using 100 μ M (Figure 3-5).

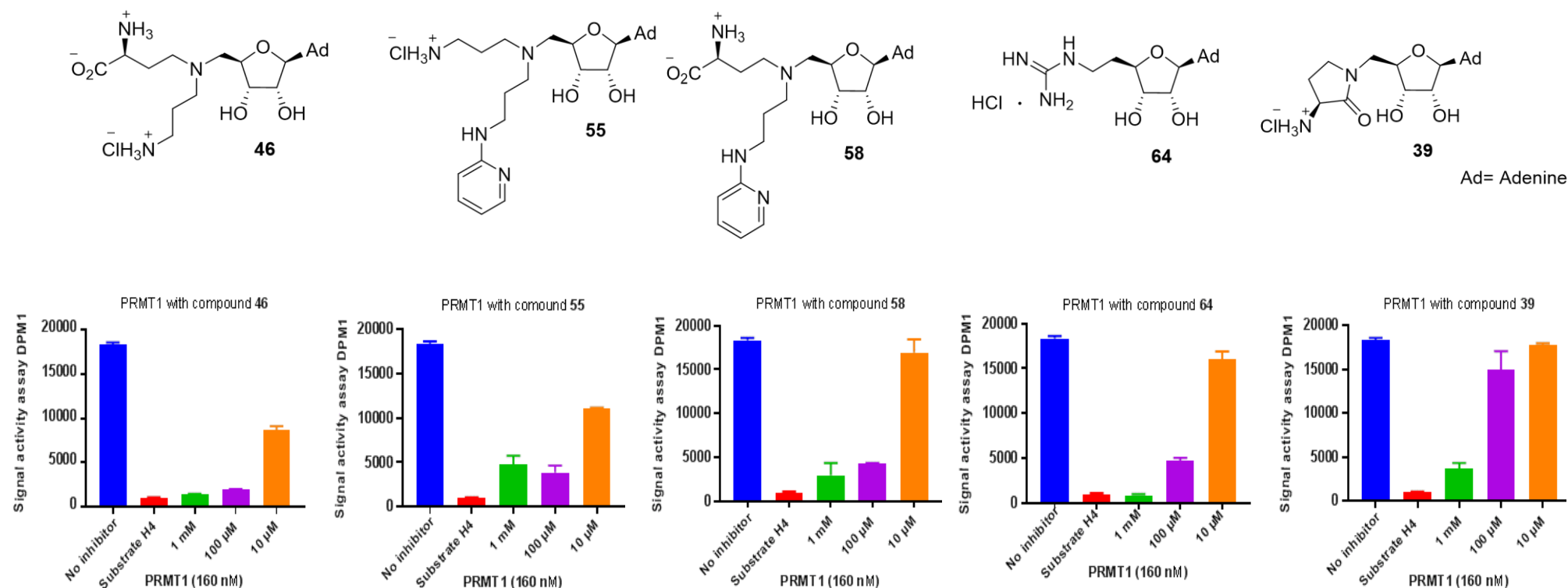


Figure 3-5: In vitro methylation assay data for compounds with PRMT1-N-His (residues 22-361).

Error bars show the standard deviation of duplicate measurements at each inhibitor concentration. Average data from in vitro methylation assay show PRMT1 (160 nM) in complex with potential small molecule inhibitors (**46**, **55**, **58**, **64**, and **39**). Two control reactions were conducted, one in the absence of inhibitor and one in the absence of PRMT1 plotted against time. Each data point corresponds to PRMT1 (160 nM) with a range of inhibitor concentrations. For the remainder of the substrate H4 1 to 21 residues is a no enzyme as a negative control.

3.2.4. Isothermal titration calorimetry (ITC) with inhibitor compounds

Obtaining a pure PRMT1 protein allowed us to study its binding with inhibitory compounds. Isothermal titration calorimetry (ITC) was used to measure the binding of various small molecule inhibitors to recombinant PRMT1. An initial objective was to titrate SAH as a model ligand into PRMT1 to establish a binding constant for comparison when using bisubstrate inhibitors. 30 μM PRMT1 in buffer 300 mM NaCl, 1 % (v/v) glycerol, 20 mM Tris-HCl and pH 7.5 was loaded into the sample cell at 25 °C. The SAH or bisubstrate inhibitors were diluted to 300 μM in the same buffer and loaded into the syringe. The binding curve obtained was fitted to a one-site model (Figure 3-6), and detailed thermodynamic parameters of binding were obtained (Table 3-1), which established that SAH bound to PRMT1 with a K_d of $1.93 \pm 0.32 \mu\text{M}$ which is comparable to CARM1 at $1.16 \pm 0.17 \mu\text{M}$.

Titration of the AzaSAM derived bisubstrate inhibitors (300 μM) against PRMT1 (30 μM) showed the propylamine derived AzaSAM **46** bound with a K_d of $6.31 \pm 0.74 \mu\text{M}$ with PRMT1; this result indicated that this compound had a binding affinity comparable to the binding of SAH to PRMT1. Surprisingly, compound **58**, where the primary amine is replaced with a 2-aminopyridine moiety designed to act as a guanidine isostere showed significantly reduced binding to PRMT1 with a K_d of $94.4 \pm 20.2 \mu\text{M}$. Interestingly, analogue **55**, which is identical to **58** but does not feature a carboxyl group, showed some binding to PRMT1 with a K_d of $4.74 \pm 1.02 \mu\text{M}$. Perhaps the removal of the carboxyl allows greater flexibility in terms of where either the 2-aminopyridine or primary amine bind to PRMT1, perhaps exchanging positions. We also evaluated compounds containing guanidinium group **64** and lactam **39**; however, no binding to PRMT1 was observed for these compounds under the conditions tested.

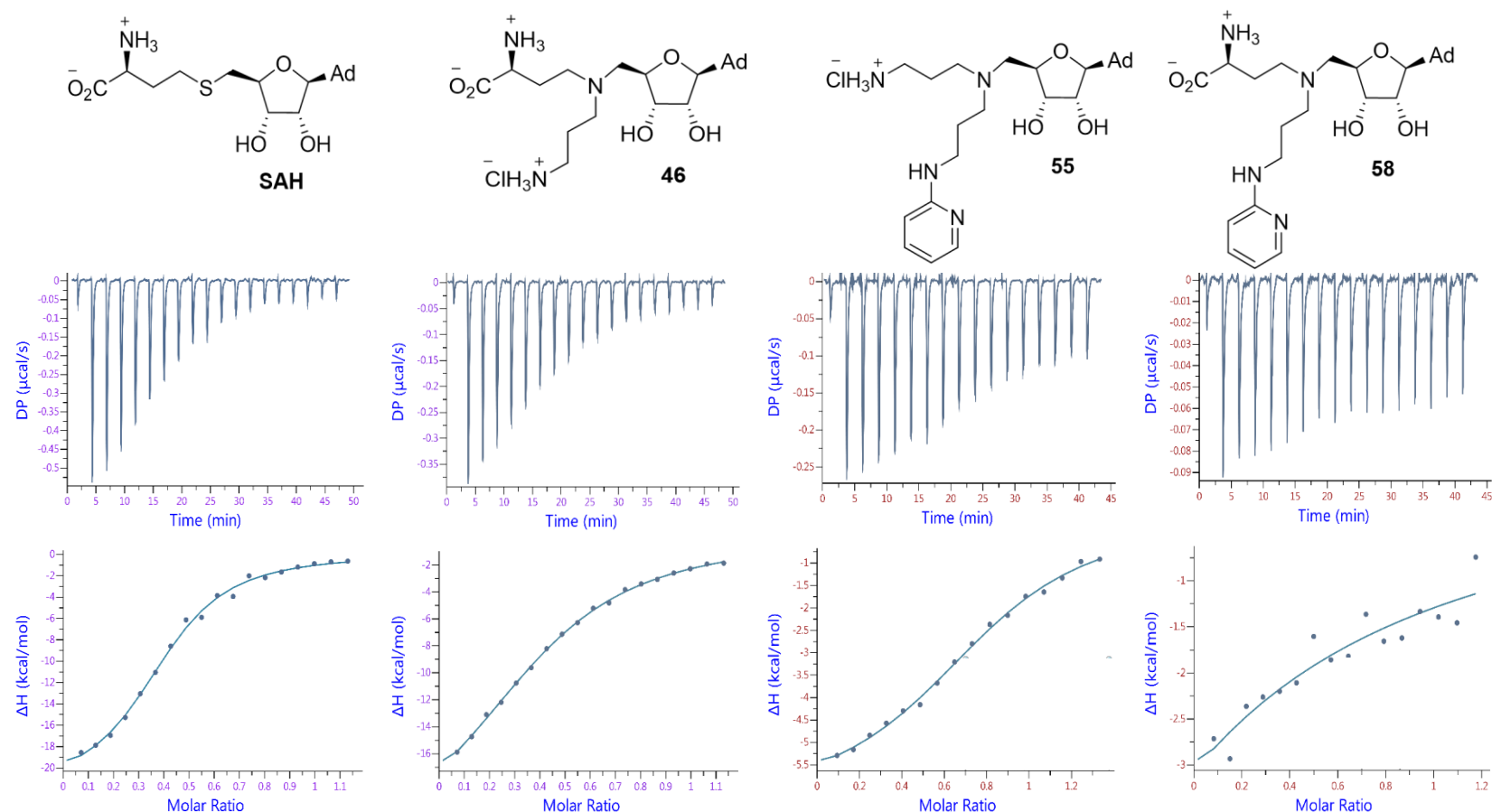


Figure 3-6: ITC data for the titration of SAH and compounds into PRMT1. The top panel represents the raw ITC data where heater power is plotted against time. Each negative spike represents an exothermic binding reaction, and as the species in the cell starts to become saturated, the spikes get smaller until only buffer-buffer spikes are observed. The bottom panel shows the binding curve produced when each spike is integrated, and energy is plotted against the molar ratio. ITC data show the binding curves for SAH and compounds **46**, **55** to PRMT1. However, bisubstrate inhibitor **58** showed no binding for PRMT1. The K_d values determined by curve fitting are given in Table 3-1 (below).

ITC results (Figure 3-6), showing binding between PRMT1 and small molecule inhibitors are summarised in Table 3-1 (data for negative controls is not shown because of poor fit to curve).

PRMT1 +ligand	Kd ± standard error (µM)	ΔG (kcal/Mol)	ΔH (kcal/Mol)	ΔS (Kcal/Mol)
SAH	1.93 ± 0.32	-7.80	-22.2 ± 1.1	15.2
46	6.31 ± 0.74	-7.09	-24.0 ± 1.4	16.9
55	4.74 ± 1.02	-7.26	-6.35 ± 0.5	10.8
58	94.4 ± 20.0	-5.49	-80.0 ± 17.0	22.5

Table 3-1: ITC-derived thermodynamic parameters for the binding of ligands **46**, **55** and **58** with PRMT1 compared to the measured affinity of SAH. The error bars for Kd values are standard errors from the curve-fitting.

3.2.5. Crystallisation assays with hPRMT1 and bisubstrate inhibitors

We attempted to crystallise PRMT1 when in complex with a selection of bisubstrate inhibitors, with the aim of gaining valuable structural information of how the inhibitors interacted with PRMT1 active site. Crystallisation trials were performed with PRMT1 (residues 22-361-6×His) construct in the presence of one of the inhibitor compounds **46**, **55** or **58** using a concentration of between 3 and 2.5 mg/mL with a 1:1 protein: compound ratio using both PACT and JCSG+ crystallisation screen⁸⁹ buffer systems. Unfortunately, only protein precipitation was observed, and no crystals were obtained after one month in these conditions.

3.3. CARM1: purification and crystallisation studies

3.3.1. Construct designed of CARM1

When crystallisation studies commenced, previous studies had reported several crystal structures of the catalytic domain of CARM1, and these were available in the protein data bank (PDB) (e.g., PDB codes 5DXJ, 4B3J, 3B3G, 3B3F, 2V7E, 2V74, 2Y1X, and 2Y1W). These crystal structures were derived from genes originating from various organisms and produced using different host species. Human CARM1 construct was produced in *E. coli*; this construct included residues S135 to T479, which represented the catalytic domain as revealed by a published crystal structure (PDB code 2Y1X).⁴¹ This CARM1 with a cleavable MBP solubility tag construct was successfully expressed and purified.⁴¹ Previously, Dr Emma Gunnell in our group had cloned MBP-CARM1-His, and biochemical studies were continued using this construct.⁸³

3.3.2. Purification of MBP-CARM1 (residues 135-479-6xHis)

The catalytic domains of CARM1 were cloned into pMAL-c2xE.⁸³ The N-terminus of CARM1 was modified with an MBP tag to increase the solubility of target proteins.^{90,91} This tag can also facilitate proper folding of recombinant proteins produced in *E. coli*.⁹⁰ The pMAL vector encoded the maltose-binding protein (MBP (41.0 kDa)) for use as an N-terminal fusion tag which can increase the solubility of target proteins. It was chosen to produce CARM1 (residues 135-479-6His) by adding 6×His tag to the C-terminus in order to allow isolation of the fusion protein from the crude lysate by the nickel affinity chromatography. A TEV cleavage site

(ENLYFQ sequence) was added between the two domains to allow subsequent removal of the MBP fusion from the expressed protein (Figure 3-7).

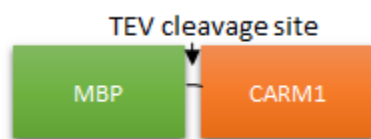


Figure 3-7: TEV protease inserted between MBP fusion tag and CARM1 (residues 135-479-6×His).

MBP-CARM1-His (residues 135-479-6×His) protein construct was produced using *E. coli* LOBSTR BL21 (DE3)-RIL cells and was initially purified by using metal ion affinity chromatography. MBP-CARM1-His typically eluted at ~25 – 35 % of buffer B, with the presence of a contaminating protein co-eluting in the first peak (Figure 3-8). A second peak eluted at ~35–45 % of buffer B, which contained pure protein without any contaminating proteins. These fractions containing pure protein were pooled and concentrated to 0.5-1 mL with a view to cleaving the MBP tag from CARM1-His by using TEV protease prior to gel filtration.

Initially, a small amount of concentrated MBP-CARM1-His protein above was used in experiments to determine how much TEV protease (~28.0 kDa) was required in order to cleave the MBP fusion protein away from the fusion construct. Assays were run by subjecting this protein to various quantities of TEV protease, and cleavage of MBP tag from target protein was analysed using SDS-PAGE (Figure 3-8B). It was found that around 0.05 molar equivalents of protease were adequate to cleave away MBP within 2-3 hours at room temperature, and these conditions were applied to the combined fractions.

After cleavage, the sample was subjected to gel filtration using a HiLoad 16/16 column (Superdex™ 200 prep grad GE Healthcare) in order to separate CARM1-His from the MBP-tag. CARM1, like other PRMTs, has been reported to form oligomers under non-reducing conditions, whereas MBP soluble tag is a stable monomer.⁹² This difference suggests that the first peak that eluted at around (65 mL, 40.0 kDa) corresponded to CARM1 (residues 135-479-6×His) which is consistent with its propensity to form oligomers or dimers (Figure 3-8C). The second peak eluted at around (82 mL, 41.0 kDa), which suggests that the MBP tag was present as a soluble monomer. Pure protein was obtained from the gel filtration as analysed by SDS-PAGE, but a total yield of CARM1 (residues 135-479-6×His) was only ~0.07 mg/L of bacterial culture. Due to the low yields obtained for this construct, alternative 2GKG soluble tags were investigated, with an aim in improving protein yield (Section 3.3.3).

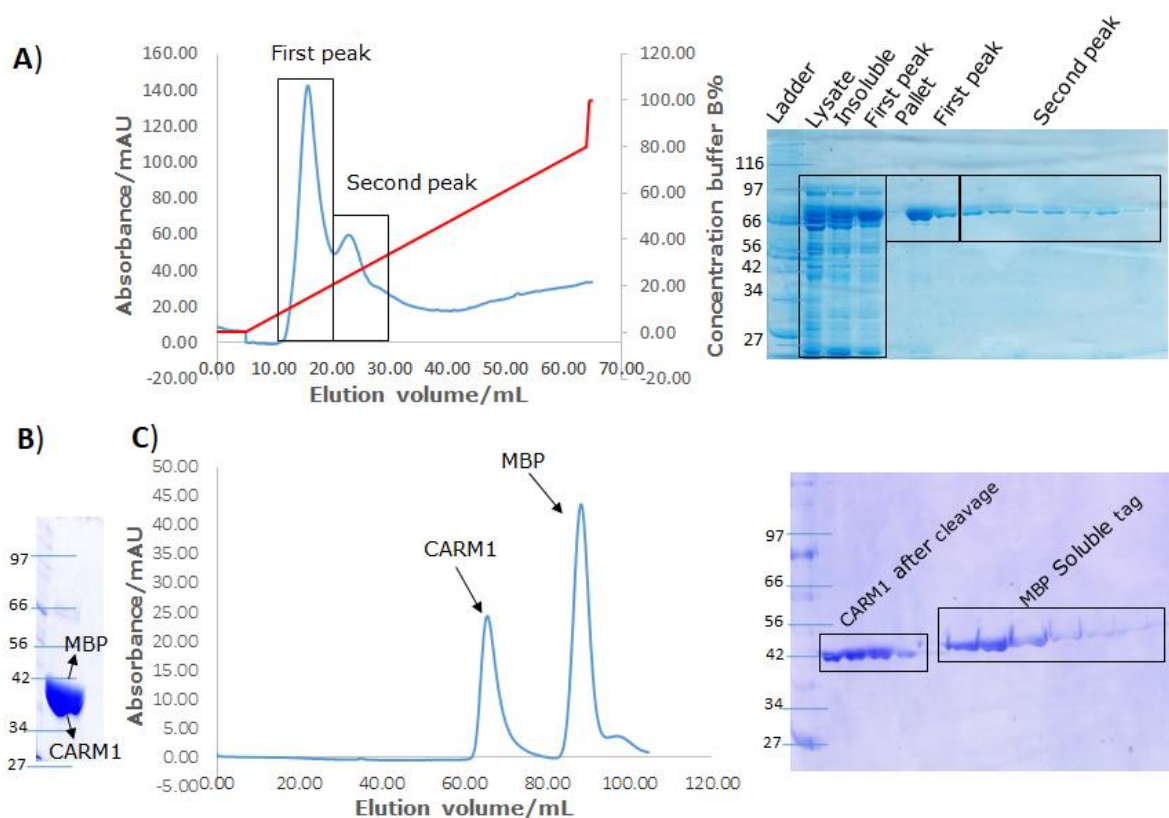


Figure 3-8: Purification of CARM1-MBP-His (residues 135-479-6xHis).

A) Nickel column purification of MBP-CARM1-His (MW= 81.0 kDa) with a corresponding SDS-PAGE gel. B) SDS-PAGE of CARM1 after MBP soluble tag removal by cleavage using TEV protease but is difficult to distinguish between them using SDS-PAGE due to nearly the similar molecular weights of CARM1 and MBP (40.0 and 41.0 kDa respectively). C) Elution profile obtained from gel filtration chromatography. SDS-PAGE gel fractions were collected from two separation peaks, which included CARM1 residues 135-479-6xHis as a dimer (MW = 39.0 kDa) and MBP tag (MW = 42.0 kDa).

3.3.3. Improving the yield of CARM1 production using an alternative tag

Protein yield needed to be improved to provide sufficient material for various experiments, such as ITC. A construct was designed to use alternative solubility tags which may assist in the folding of fused proteins and therefore, increase protein yield. The previous MBP soluble tag was replaced with a 2GKG tag⁹³ in the

pRSF-13 vector, previously identified by Dreveny's group (unpublished results).⁹⁴ This tag was derived from a receiver domain originally isolated from a *Myxococcus xanthus* social motility protein FrzS (PDB code 2GKG).⁹³

Initially, two CARM1 constructs were cloned: one with a cleavable 2GKG soluble tag at the N-terminus and a 6×His tag at the C-terminus (2GKG-CARM1-His) (Figure 3-9). The second contained a non-cleavable 2GKG soluble tag at the N-terminus and a 6×His tag at the C-terminus (2GKG-CARM1-His). The construct with the non-cleavable tag will be discussed in more detail later in this chapter (see Section 3.4.).

3.3.4. CARM1 with a cleavable 2GKG soluble tag

CARM1 with a cleavable tag (residues 135-479-6×His) was cloned into a pRSF-13 vector; as a 2GKG fusion protein, this protein also contained a TEV cleavage site for subsequent cleavage and removal of the 2GKG tag. The plasmid was transformed into LOBSTR BL21 (DE3)-RIL cells for expression; production and purification of 2GKG-CARM1-His were carried out in the same manner as described for MBP-CARM-His (Figure 3-8). Protein purification was initially achieved by loading the produced protein onto the nickel affinity chromatography column. Two peaks were obtained from the nickel column; the first was a small peak that eluted with 20 % of buffer B, which contained contaminating protein. The second was a more massive, broader peak at 50 % of buffer B; this peak appeared to be 2GKG-CARM1-His after analysis by SDS-PAGE.

Cleavage of CARM1 from the 2GKG soluble tag was performed using TEV protease using the same ratio, as described above (see Section 3.3.2.). The cleaved protein was then subjected to analysis on a gel filtration Superdex™ 200 HiLoad 16/16 column (GE Healthcare). Three peaks were observed following the gel filtration, indicating that CARM1 eluted at 65 mL, as expected for oligomeric proteins typically seen for PRMTs, whereas both TEV protease and 2GKG soluble tag eluted separately as monomers at 85 and 95 mL, respectively. Examination of these fractions by SDS-PAGE gel revealed bands, which corresponded to the molecular weight of CARM1 at approximately 40.0 kDa, TEV 27.0 kDa and 2GKG 14.0 kDa, respectively (Figure 3-9).

The yield of purified CARM1 protein obtained was 0.24 mg/L of bacterial culture, which was higher than that observed for the purified protein using an MBP tag (0.07 mg/L). This improvement reflected the increased proportion of contaminant-free nickel fractions, which could be taken forward for subsequent purification steps; rather than optimisation of bacterial expression or an improvement in protein solubility.

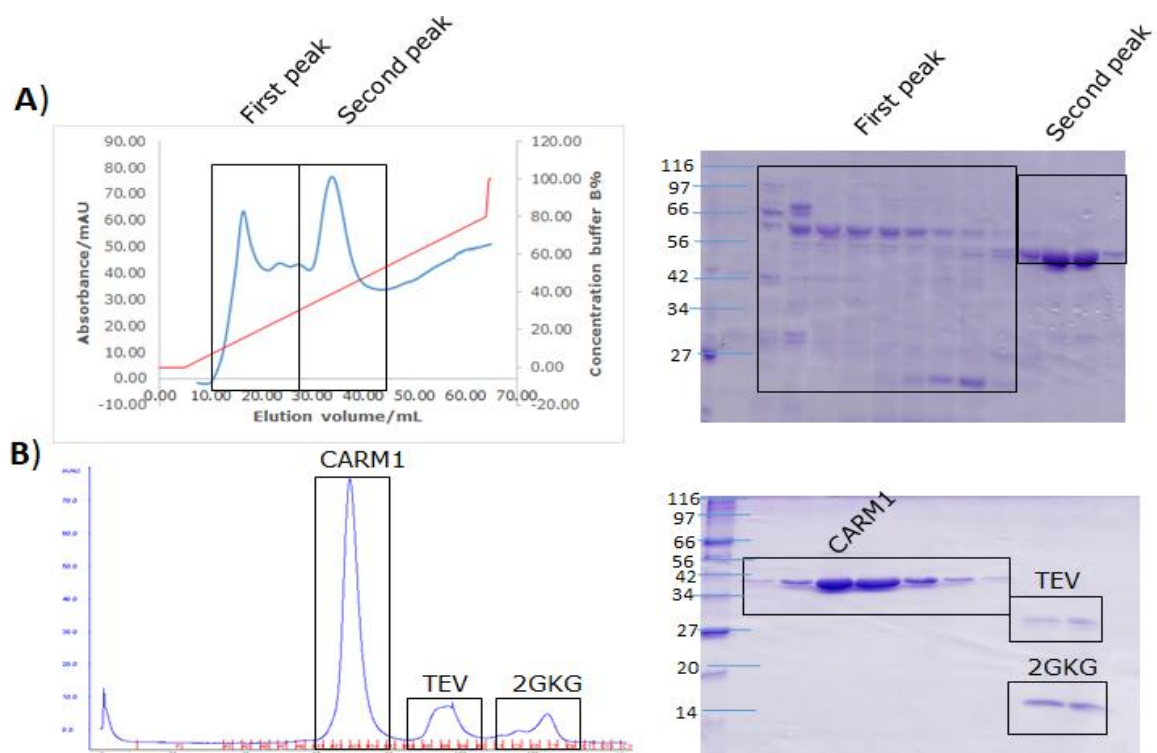


Figure 3-9: Purification of CARM1 construct with cleavable 2GKG soluble tag.

A) Chromatogram of the nickel column with a corresponding SDS-PAGE analysis of the fractions. B) UV trace of the gel filtration showed three main peaks indicating CARM1 (residues 135-479-6×His) (MW=39.0 kDa), TEV protease (MW=27.0 kDa) and the 2GKG tag (MW=14.0 kDa). SDS-PAGE analysis containing fractions from the gel filtration affinity, which were combined for gel filtration, are boxed.

3.3.5. *In vitro* methylation assay to test the activity of compounds with CARM1

An *in vitro* methylation assay was used to compare CARM1 inhibition when incubated with selected compounds. The experiment was performed using CARM1 (230 nM), target substrate H3 (residues 3-31-6×His, 120 μM), and [³H]-SAM 0.23 μM, in the presence of a range of inhibitors **46**, **55**, **58**, **64**, and **39** at various concentrations (1.0 mM, 100 μM and 10 μM). The experiments were performed by Dr Emma Gunnell. The four bisubstrate inhibitors **46**, **55**, **58** and **64** displayed

robust inhibition of CARM1 activity at 10 μ M. Addition of these inhibitors at this concentration reduced CARM1 activity by 95 % when compared to activities observed under control conditions (Figure 3-10). Lactam **39** showed weak inhibition activity at this concentration. Assays using 100 μ M revealed that lactam reduced CARM1 activity by 15 % when compared with activity observed under control conditions.

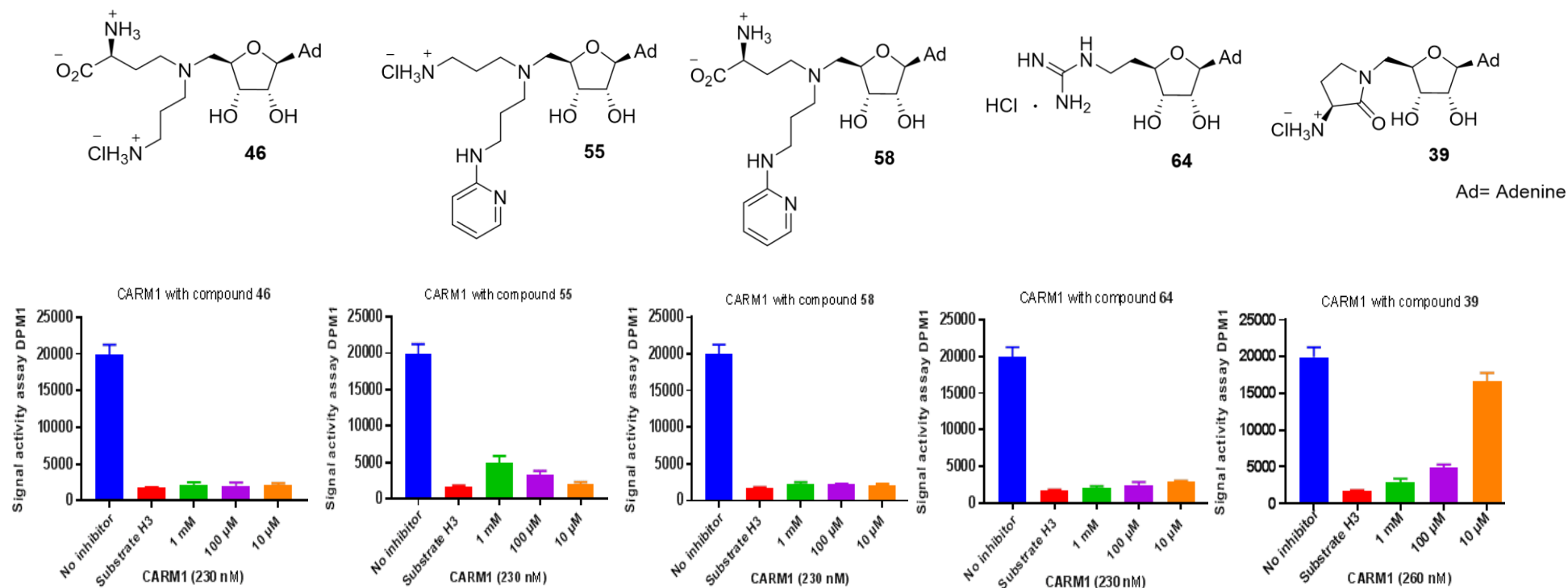


Figure 3-10: Bisubstrate inhibitors evaluation against CARM1-N-His. Two control reactions were shown, one in the presence of CARM1 and substrate H3 (3-31 residues) in the presence of [^{3} H]-SAM 0.23 μ M in blue as a positive control, and one in the absence of CARM1 in red as a negative control. Each set of other data points are labelled, and each corresponds to CARM1 (260 nM) with the substrate (120 μ M) and [^{3} H]-SAM 0.23 μ M in the presence of different concentrations of compounds **46**, **55**, **58**, **64**, and **39**. Error bars show the standard deviation of duplicate measurements at each inhibitor concentration. For the remainder of the substrate H3 residues 3 to 31 residues is a no enzyme as a negative control

3.3.6. Isothermal titration calorimetry (ITC) and IC₅₀ for CARM1

The selective binding of compounds **46**, **55**, **58**, **64**, and **39** to CARM1 was investigated using the ITC method; this method had been successfully used when investigating binding affinities of PRMT1 (see Section 3.2.4.). In general, CARM1 concentrations used in the assays were between 20 and 30 μM , while the concentrations of inhibitory compounds **46**, **55**, **58**, **64**, and **39** were 200 –300 μM , respectively. Initially, assays with SAH titrated onto CARM1, resulted in a K_d of $1.16 \pm 0.17 \mu\text{M}$ which compared with data from literature which shows SAH binding to CARM1 with a K_d of $6.90 \mu\text{M}$.⁴¹

Assays with the isopropylamine derived general bisubstrate inhibitor **46** revealed that this compound bound more tightly to CARM1 with a K_d $0.59 \pm 0.16 \mu\text{M}$ which compared to PRMT1 where the K_d was $6.31 \pm 0.74 \mu\text{M}$ (Section 3.2.4). Interestingly, the 2-aminopyridine derived bisubstrate inhibitor **55** also showed strong binding to CARM1 with a K_d $1.15 \pm 0.40 \mu\text{M}$; this was slightly stronger than that observed for binding to PRMT1 with K_d $4.74 \pm 1.0 \mu\text{M}$. In comparison to PRMT1, there seemed to be no penalty for removing the carboxyl group in terms of binding to CARM1, since compound **58** had a K_d of $1.07 \pm 0.12 \mu\text{M}$ when bound to CARM1, with a similar dissociation constant as that observed for **55** with a K_d $1.15 \pm 0.40 \mu\text{M}$ (Figure 3-11). These results may indicate a significant preference for the binding of compounds **46**, **55**, and **58** to CARM1 when compared with PRMT1.

These compounds showed an approximate ~15- and ~40-fold selectivity towards CARM1 when compared with PRMT1. These are preliminary results, but the potency of the 2-aminopyridine (**55** and **58**) derived inhibitors of CARM1

encouraged further investigation (*e.g.*, confirmed selectivity over other methyltransferases, higher potency and demonstrated cellular activity). These criteria would need to be met before bisubstrate inhibitors could be used to provide biological information in regards CARM1 inhibition.

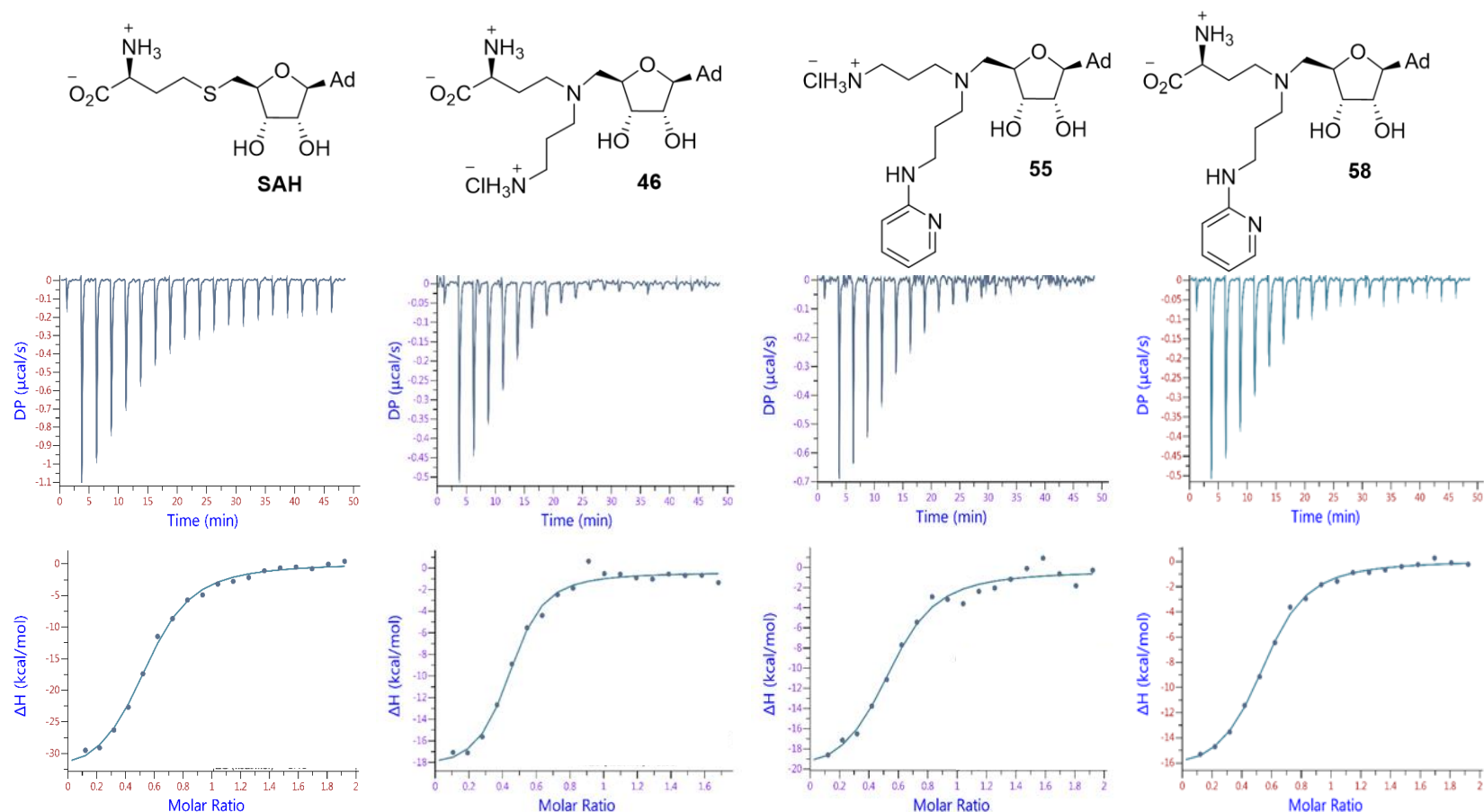


Figure 3-11: ITC data for the titration of bisubstrate inhibitors (200 μM) into human CARM1 (residues 135-479-6 $\times$ His) (20 μM). The error bars for K_d values are standard errors from curve-fitting. SAH was used as a positive control when binding to CARM1 and curves for titration of **46**, **55** and **58** to CARM1 are shown. In each case, the top panel is the raw ITC data showing the exothermic release of heat upon bisubstrate inhibitor binding, and the bottom panel is the isotherm produced by the integration of the peaks. The K_d values determined by curve fitting are given in Table 3-2.

In addition to K_d values, the thermodynamic parameters arising from inhibitor binding was also determined (Table 3-2). This data revealed that binding for all ligands tested was enthalpically favourable (*i.e.*, $\Delta H < 0$) and entropically unfavourable ($-\Delta S > 0$). This event is consistent with multiple electrostatic interactions being involved in binding with the compounds overcoming the flexibility of the ligands.

CARM1 + ligand	$K_d \pm$ standard error (μM)	ΔG (kcal/Mol)	ΔH (kcal/Mol)	ΔS (Kcal/Mol)
CARM1+SAH	1.16 ± 0.17	-8.10	-34.6 ± 1.2	11.0
CARM1+ 46	0.59 ± 0.16	-8.50	-18.5 ± 0.9	9.96
CARM1+ 55	1.15 ± 0.40	-8.11	-20.7 ± 1.7	12.6
CARM1+ 58	1.07 ± 0.12	-8.15	-17.2 ± 0.4	9.01

Table 3-2: ITC-derived thermodynamic properties for the binding of ligands with CARM1. The table shows a summary of the thermodynamic properties when either bisubstrate inhibitor or SAH when titrated into a solution of human CARM1 (residues 135-479-6×His) as indicated. The error bars for K_d values are standard errors from curve-fitting.

The inhibition binding of PRMT1 and CARM1 and IC_{50} were compared (Table 3-3) (IC_{50} values measured by Dr Emma Gunnell). K_d values were determined within the expected range, based on previously reported K_d and IC_{50} values.^{95,96} Increased binding of compounds (**55** and **58**) containing a 2-aminopyridine group was observed for CARM1 over PRMT1, as suggested by IC_{50} values arising from the methylation assay. This effect was not apparent for PRMT1 due to low overall potency against this protein.

Compound number	K _d ± standard error (μM)		IC ₅₀ (95 % confidence interval) (μM)	
	PRMT1	CARM1	PRMT1	CARM1
SAH	1.93 ± 0.32	1.16 ± 0.17	0.5 (0.2 - 1.0)	0.2 (0.1 - 0.4)
46	6.31 ± 0.74	0.59 ± 0.16	-	-
55	4.74 ± 1.02	1.15 ± 0.40	52 (36 - 74)	0.18 (0.13 - 0.25)
58	94.4 ± 20.3	1.07 ± 0.12	9.7 (7.2 - 13)	0.75 (0.68 - 0.82)

Table 3-3: K_d and IC₅₀ values for our bisubstrate inhibitors against PRMT1 and CARM1.

Summary of K_d and IC₅₀ values for compounds with PRMT1 and CARM1 measured using ITC and enzyme activity assays, respectively. Dashes indicate that values have not been determined.

3.3.6.1. Crystallisation of CARM1 with small molecule inhibitors

With pure protein required for the best possibility of obtaining crystals, attention turned to attempting to co-crystallise the catalytic domain of CARM1 (residues 135-479-6His) with the compounds containing the 2-aminopyridine group **55** and **58**, bisubstrate inhibitor **46** and lactam compound **39**, respectively.

Successful crystallisation of a protein typically requires concentrations between 1-50 mg/mL of protein depending on its solubility.⁹⁷ Previous reports indicated the concentration required for CARM1 to be in the range of 1-2.5 mg/mL.^{41,83} Crystallisation trials were thus performed with CARM1 at 1.30 mg/mL or 32 μM in complex with these molecules. A PACT screen was used and drops set either with or without 320 μM of inhibitors. Crystallisation trials were incubated at 19 °C and used a 96-well screen (PACT premier). After one week, several crystals were obtained, a selection of which are shown in (Table 3-4) and (Figure 3-12).

Interestingly, no crystals were observed in the absence of the small molecules and precipitation was observed with the majority of drops. This indicated that the presence of the bisubstrate inhibitors promoted crystal formation, suggesting that the bisubstrate inhibitors were binding to the active site of CARM1, and their presence was enhancing the stability of the protein.

PACT Screen no	Premier	Concentration of Salt	Buffer	pH	Precipitant
C7		0.2 M sodium chloride	0.1 M HEPES	7.0	20 % w/v PEG 6000
D8		0.2 M ammonium chloride	0.1 M Tris	8.0	20 % w/v PEG 6000
D10		0.2 M magnesium chloride hexahydrate	0.1 M Tris	8.0	20 % w/v PEG 6000
E7		0.2 M sodium acetate trihydrate			20 % w/v PEG 3350
E12		0.2 M sodium malonate dibasic monohydrate			20 % w/v PEG 3350
F6		0.2 M sodium formate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
F7		0.2 M sodium acetate trihydrate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
F8		0.2 M sodium sulfate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
F10		0.02 M sodium/potassium phosphate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
F11		0.2 M sodium citrate tribasic dihydrate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
F12		0.2 M sodium malonate dibasic monohydrate	0.1 M Bis-Tris propane	6.5	20 % w/v PEG 3350
G5		0.2 M sodium fluoride	0.1 M Bis-Tris propane	7.5	20 % w/v PEG 3350
H1		0.2 M sodium fluoride	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350
H2		0.2 M sodium bromide	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350
H6		0.2 M sodium formate	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350
H7		0.2 M sodium acetate trihydrate	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350
H11		0.2 M sodium citrate tribasic dihydrate	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350
H12		0.2 M sodium malonate dibasic monohydrate	0.1 M Bis-Tris propane	8.5	20 % w/v PEG 3350

Table 3-4: Crystallisation conditions used to generate CARM1-inhibitor co-crystals. Summary of the needle crystals that were obtained from PACT premier screens that were used to crystallise CARM1 when complexed with different bisubstrate inhibitors **46**, **55**, **58** and **39**. All crystals were grown at 19 °C and formed within three days.

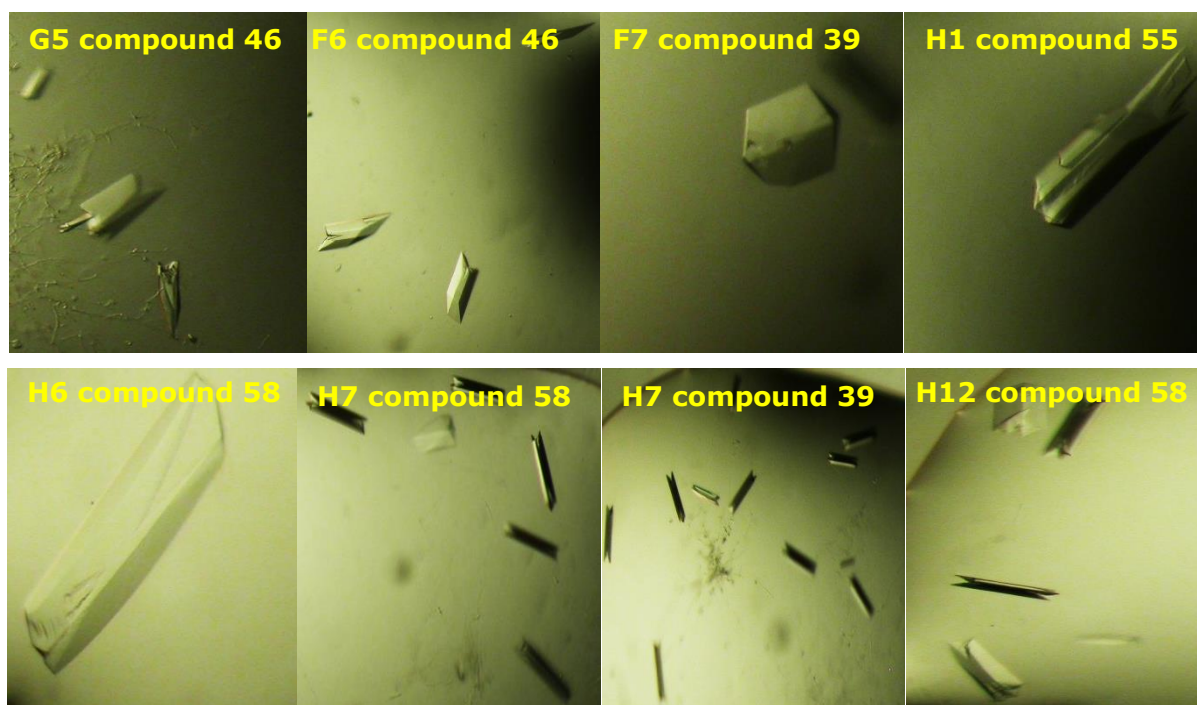


Figure 3-12: Crystal obtained from crystallisation condition PACT premier. CARM1 crystals when complexed with compounds **46**, **55**, **58** and **39** as indicated, in different crystallisation conditions using a protein concentration of 3 mg/mL. Codes referring to PACT conditions are detailed in the experimental section. Crystals that diffracted to the highest resolution were obtained from H1 and H6 crystallisation conditions (H1 0.2 M NaF 0.1 M Bis-tris propane 8.5 20 % w/v PEG 3350) and (H6 0.2 M sodium formate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350) with the compounds **46**, **55**, **58** and **39**, respectively.

3.3.6.2. Data collection and structure determination

A selection of crystals was successfully flash cooled in liquid nitrogen and taken to the Diamond Light Source (beamline I04). A sample image of the cryo-cooled crystal in a loop and diffraction image are shown (Figure 3-13).

Between 1600-1800 diffraction images with an oscillation angle of 0.1° (Table 3-5) were obtained. Sample crystals of CARM1 in complex with various inhibitors (shown in brackets) arising from PACT Primer H1, H6 and H12 conditions diffracted

to resolutions of 2.44 Å (**55**), 2.01 Å (**58**) and 2.33 Å (**46**) respectively and generally displayed the same space group P 2₁ 2₁ 2 and unit cell parameters (Table 3-5). Frequently, the maximum diffraction resolution correlated with crystal morphology; a CARM1 crystal in complex with compound **39** gave a weak X-Ray diffraction with a resolution of 3.3 Å, which was too low to solve the structure. The low resolution for this compound **39** in complex with the catalytic domain of CARM1 is a common feature for many protein crystals. This could be high solvent volume within the crystal or may be due to poor order of molecules.¹¹² There are a number of techniques which can be implemented to improve diffraction resolutions, including crystal dehydration, which is used to reduce the solvent volume and improve crystal packing;¹¹³ and seeding, which involves using a crystal as a nucleation point for the growth of larger crystals.¹¹⁴ These techniques can be carried out in future crystallisation trials with the aim of increasing diffraction resolutions.

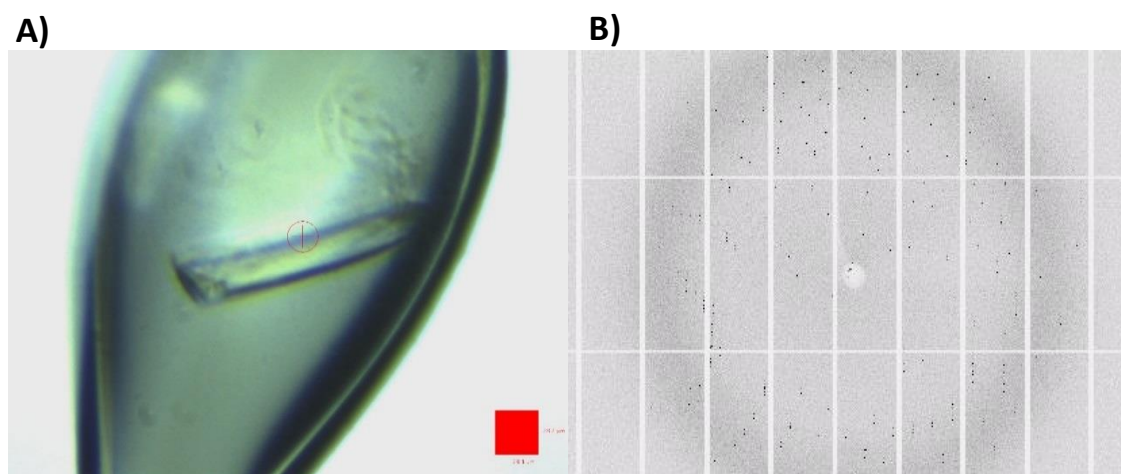


Figure 3-13: Sample diffraction image from the crystal of CARM1 with compound **55**.

A) Large crystals of CARM1 in complex with compound **55** from PACT screen crystallisation (H6) were mounted in a loop cryo-cooled. B) Sample diffraction image collected at the Diamond Light Source from a CARM1 crystal with a small molecule inhibitor that diffracted to the highest resolution obtained. This crystal was obtained from the optimised H6 PACT screen crystallisation condition (0.2 M sodium formate, 0.1 M bis-tris propane, pH 8.5, 20 % w/v PEG 3350).

Data set	Compound 58	Compound 55	Compound 46	Compound 39
Space group	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2
a, b, c	75.51 Å 99.23 Å 208.5 Å	74.78 Å 98.62 Å 207.1 Å	75.23 Å 98.95 Å 207.9 Å	74.99 Å 99.56 Å 208.21 Å
α, β, γ	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Low-resolution limit (Å)	69.48	71.42	70.68	99.42
High-resolution limit (Å)	2.17	1.92	2.11	3.29
R _{merge}	0.779	0.644	0.679	0.938
R _{pim}	0.400	0.314	0.279	0.392
CC 1/2	0.7	0.811	0.7	0.659
No of total reflections	23005	49432	35459	3743
No of unique reflections	3884	7387	5241	547
Mean I/σI	1.9	2.50	1.5	2.0
Completeness (%)	99.2	98.2	94.4	95.7
Multiplicity	6.6	6.7	6.5	6.4

Table 3-5: Data collection statistics of data sets obtained.

Unit cell parameters, space group and data collection statistics from data sets of co-crystals of CARM1 in complex with four different small molecule inhibitors are described in the table.

The structures of CARM1 in complex with bisubstrate inhibitors (**55** and **58**) were resolved within our group by Dr Emma Gunnell, and the data for inhibitor **46** will potentially be solved in the future. Molecular replacement (MR) was used to resolve the structure of CARM1 with compounds **55** and **58**. In the initial stage, human CARM1 structure (PDB code 2Y1X),⁴⁰ was prepared as a model for MR by

removing water molecules and ligands. The resulting electron density maps (2Fo-Fc and Fo-Fc) and PDB coordinate file were opened in Coot⁹⁸ and inspected. Several rounds of restraint refinement were performed in PHENIX⁹⁹ to improve the side chain residues fits in the electron density map and manually adjusted where needed. Finally, refined structures were visualised in PyMOL¹³⁷ to identify electrostatic interaction with key residues in the active site.

3.3.6.3. Bisubstrate inhibitors binding to the active site of CARM1

Analysis of the electron density maps analogous to the active site of CARM1 confirmed that both inhibitors (**55** and **58**) were bound to the SAM-binding site and the substrate-binding channel. Analysis revealed that these compounds were within range for optimal interactions with many residues present in the active site (Figures 3-14 and 3-15).

Superposition of CARM1 in complex with compound **58** revealed a conformation where the aromatic group occupied the substrate-binding channel, and the terminal α -amino acid group (which mimics that of SAM) occupied the SAM amino acid pocket. Within the SAM-binding site, residues Arg-169 and Glu-258 interacted with the carboxylate group of compounds **58**. Ribose 2' and 3' hydroxyls formed hydrogen bonds with Glu-215. The adenine ring nitrogen N6, N1 and N7 interacted with Glu-244, Ser-272 and the main-chain nitrogen atom of Val-243, respectively. The 2-aminopyridine group showed good interactions with residues in the substrate arginine-binding cavity, especially Glu-258 with 2-aminopyridine and Glu-267 with the nitrogen of pyridine at ~ 3.2 and $2.3 \text{ \AA}$, respectively.

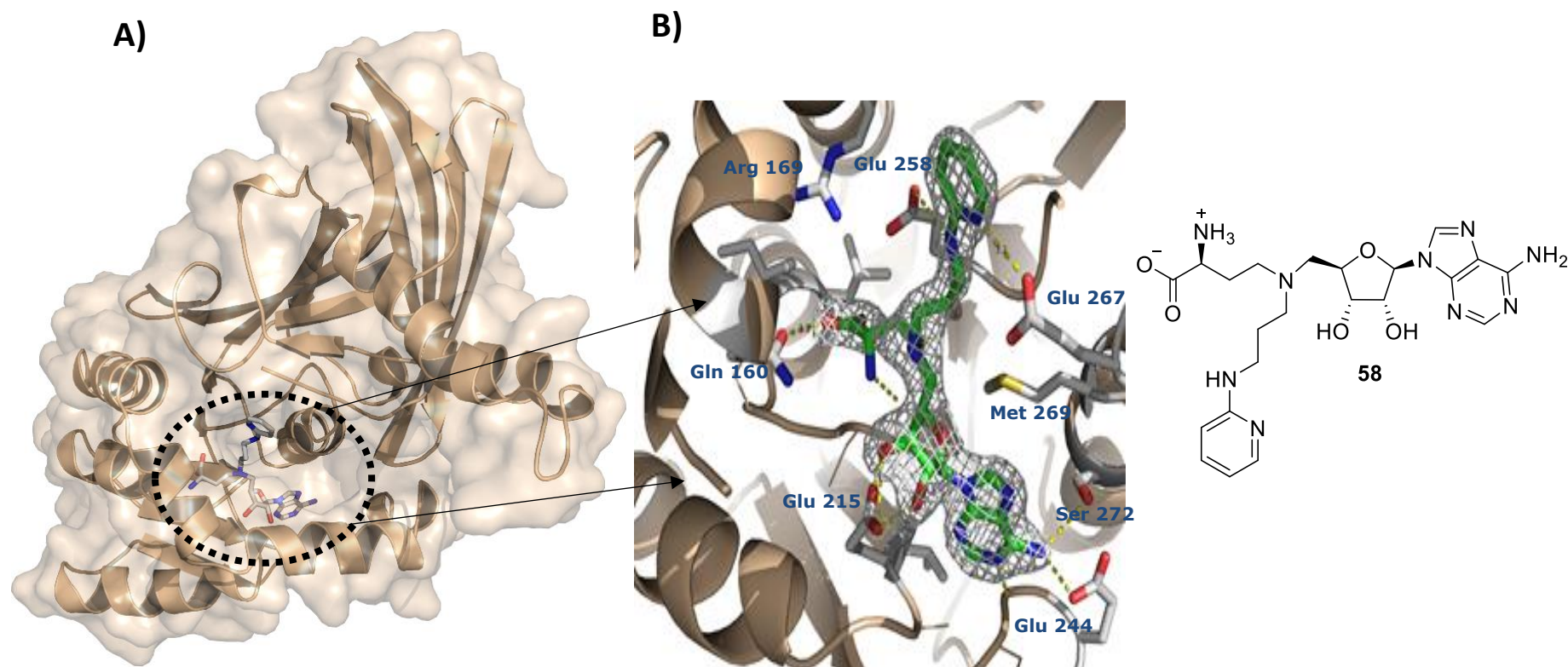


Figure 3-14: CARM1 co-crystal structures with compound **58**.

A) Superposition of compound **58** in each monomer within the asymmetric unit (wheat) of respective co-crystal structures with CARM1. B) Binding pockets of CARM1 shows compound **58** occupying both positions of SAM amino acid and substrate binding channel (highlighted as a mesh). 2mF_o-DF_c map of monomer in which density is visible for all of compound **58**. Maps are contoured at 1σ. Polar contacts found using PyMol molecular graphics system (version 2.1.0)¹³⁷ are shown as yellow dotted lines, and relevant residues involved in the SAM-binding site and arginine residue of the substrate are shown as sticks.

Examination of electron density maps for compound **55**, suggests that this bisubstrate inhibitor may adopt two conformations within the active site of CARM1, in which the 3-aminopropyl and 2-aminopyridine groups alternately occupy the SAM amino acid and binding channels of the substrate arginine, respectively (Figure 3-15). In one conformation, the 3-aminopropyl group bound at the SAM-binding site formed hydrogen bonds with Glu-258, Asp-191, and Gln-160, respectively, *via* a water molecule. Meanwhile, the 2-aminopyridine occupied the arginine substrate channel with the pyridine nitrogen, hydrogen bonding to Arg-169 and exocyclic amine mainly interacting with Glu-267, and Ser-195 *via* water. In the second conformation, the 2-aminopyridine residue bound at the SAM-binding site with the pyridine nitrogen and underwent hydrogen bonding with Gln-160 and Ser-196, respectively, *via* a water molecule. While, the 3-aminopropyl group bound at the substrate-binding site *via* Glu-258, also formed hydrogen bonds with Trp-416, and Glu-258, respectively, *via* a water molecule.

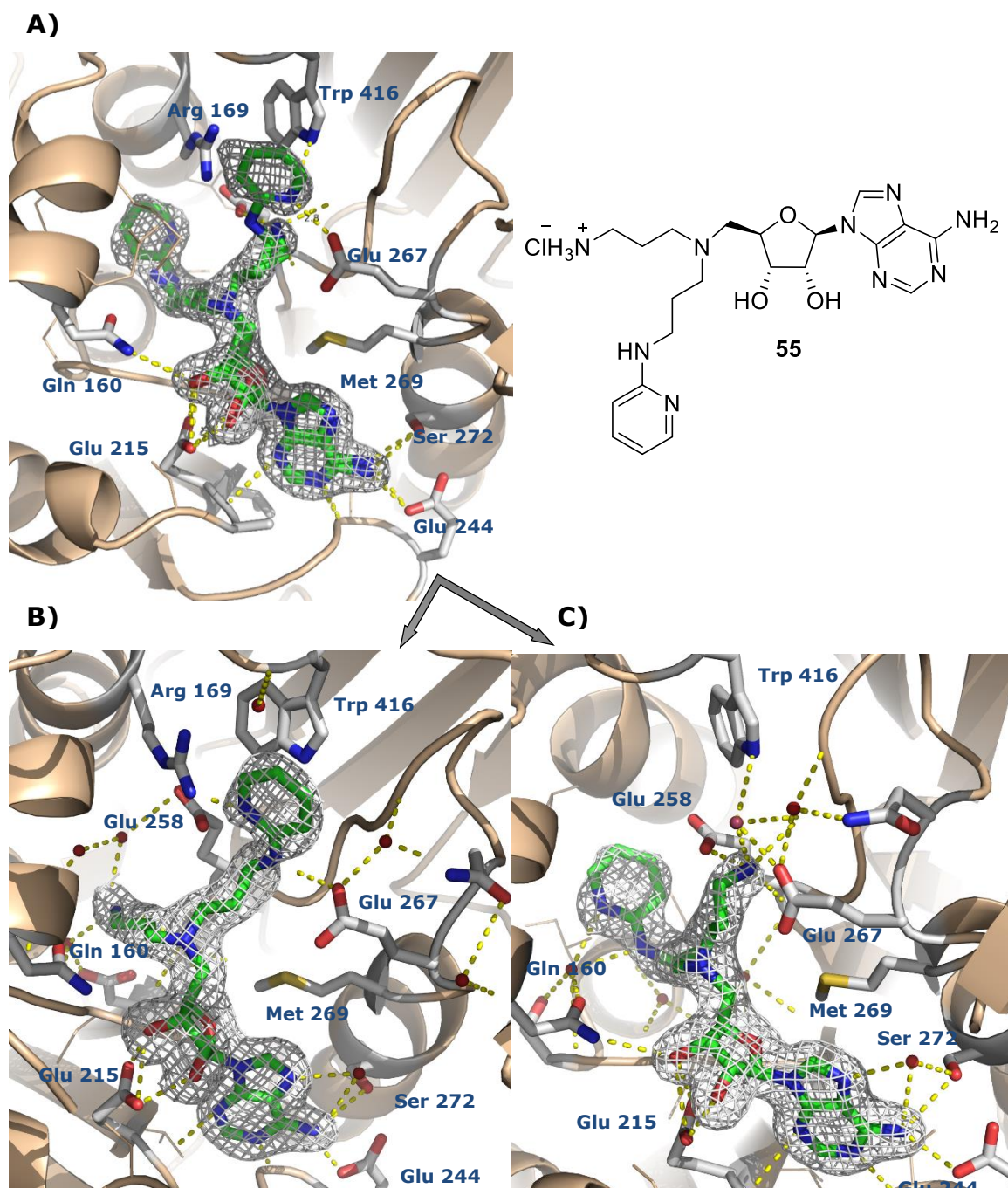


Figure 3-15: Superposition of CARM1 in complex with compound **55**.

A) Crystal structures of compound **55** carrying 2-aminopyridine and 3-aminopropyl bound to CARM1 revealed two conformations within the active site and maps are contoured at 1σ . B) Binding pockets of CARM1 show the 2-aminopyridine occupying the SAM amino acid binding pocket and 3-aminopropyl bound at the substrate arginine binding (highlighted as a mesh) C) A second conformation bound in the active site of CARM1 is shown when the 3-aminopropyl bound to SAM amino acid binding pocket and 2-aminopyridine occupying the substrate-binding site (highlighted as a mesh). Polar contacts found using PyMol¹³⁷ are shown as yellow dotted lines, and relevant residues involved in SAM binding site and the arginine residue of the substrate are shown as a stick (highlighted as a mesh).

Overall, the binding studies showed that bisubstrate inhibitors containing 2-aminopyridine **55** and **58** conferred a 15 to 40-fold selectivity for CARM1 over PRMT1. Crystal structures showed that compound **55** containing 3-aminopropyl and 2-aminopyridine groups was able to adopt alternate conformational positioning in the binding pocket of the SAM-amino acid and arginine-binding channel. Compound **58** that includes a terminal α -amino acid group adopts a single conformation, with the aminopyridine being present in the binding channels of the substrate (Figure 3-13.). The evidence suggests that the carboxylate group in compound **58** may be needed to ensure the arginine-binding channel occupies the desired pocket. The presence of a carboxylate group in compound **58** conferred a conformational preference for binding with the active site of CARM1, but a carboxylate group does not appear to enhance CARM1 affinity with other compounds.⁸³

3.4. CARM1 with a non-cleavable 2GKG soluble tag

CARM1 with a non-cleavable 2GKG tag construct was designed and cloned to gain further understanding about whether this fusion protein would fold correctly and potentially be used for comparison of binding and structure of complexes with inhibitors. This development could result in an improved overall yield of protein as fewer purification steps would be required. It was also interesting to learn exactly how the presence of a soluble tag would affect the active site position of CARM1. Binding studies using this fusion protein conducted using ITC revealed that there was no substantial effect on the active site or, presumably, how the protein folded (Figure 3.17). 2GKG-CARM1 construct was also used in crystallisation which could

provide valuable structural information of a 2GKG soluble tag linked to the catalytic domain of CARM1.

CARM1 construct (residues 135-479) with a non-cleavable 2GKG soluble tag at the N-terminus and 6×His at C-terminus was produced in *E. coli* and then purified by nickel affinity chromatography followed by gel filtration. When 2GKG-CARM1-His elution sample was applied to the nickel column, two peaks were obtained and assessed by SDS-PAGE. The first peak was characteristic of the presence of a contaminating protein co-eluting (~ 62.0 kDa) in an early fraction, whereas the second peak which is 2GKG-CARM1-His (56.0 kDa) eluted in the later fraction following assessment by SDS-PAGE. Only fractions without contamination from the second peak were combined and used for gel filtration. Analytical gel filtration (HiLoad 16/60 Superdex 200 prep grade) column showed that a small earlier eluting peak consisting of aggregated protein eluted at 45 mL, whereas the latter peak eluting at 65 mL contained proportionally more of the target protein (2GKG-CARM1-His, 56.0 kDa, Figure 3-16). Pure protein was obtained as revealed by analysis on SDS-PAGE (56.0 kDa) (Figure 3-16). The total yield obtained was 0.56 mg/L of bacterial culture which compared favourably to the yield of 0.077 mg/L culture obtained when using the MBP fusion tag which gave and was suitable for crystallisation and ITC binding studies. Following crystallisation trials, crystals formed in three conditions in 2GKG-CARM1-His trials but, at best, very poor diffraction was observed with one crystal (Section 3.4.2.).

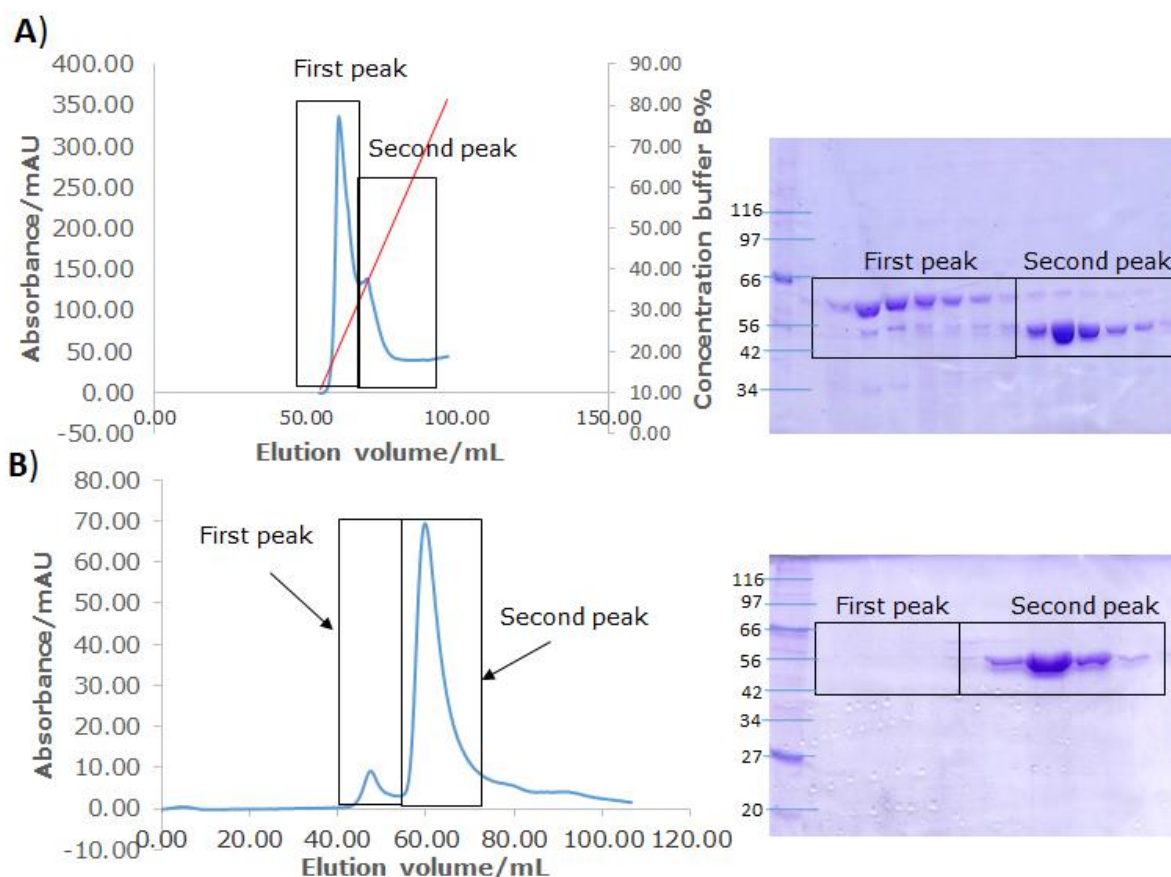


Figure 3-16: Purification of CARM1 with non-cleavable 2GKG soluble tag construct.

A) UV trace of the nickel column alongside the corresponding SDS-PAGE analysis of the fractions. B) UV trace of gel filtration (HiLoad 16/60 Superdex 200 prep grade) column exhibited of CARM1 (residues 135-479--6×His) with 2GKG soluble tag (MW = 56.0 kDa) and subsequent SDS-PAGE analysis.

3.4.1. ITC analysis for CARM1 with non-cleavable 2GKG soluble tag

ITC was used to confirm the SAH binding properties of CARM1 with a non-cleavable 2GKG soluble tag. The main aim of the experiment was to investigate whether the N-terminus 2GKG soluble tag had an effect on the active site of CARM1. ITC was performed, and the resulting binding curve fit a one-site model (Figure 3-17), to generate thermodynamic binding data (Table 3-6). ITC results revealed that SAH model ligand had bound to CARM1 with a K_d of $1.17 \pm 0.12 \mu\text{M}$, which compared

to a K_d of $1.16 \pm 0.17 \mu\text{M}$ for CARM1 without soluble tag indicating that the presence of the soluble tag did not affect binding.

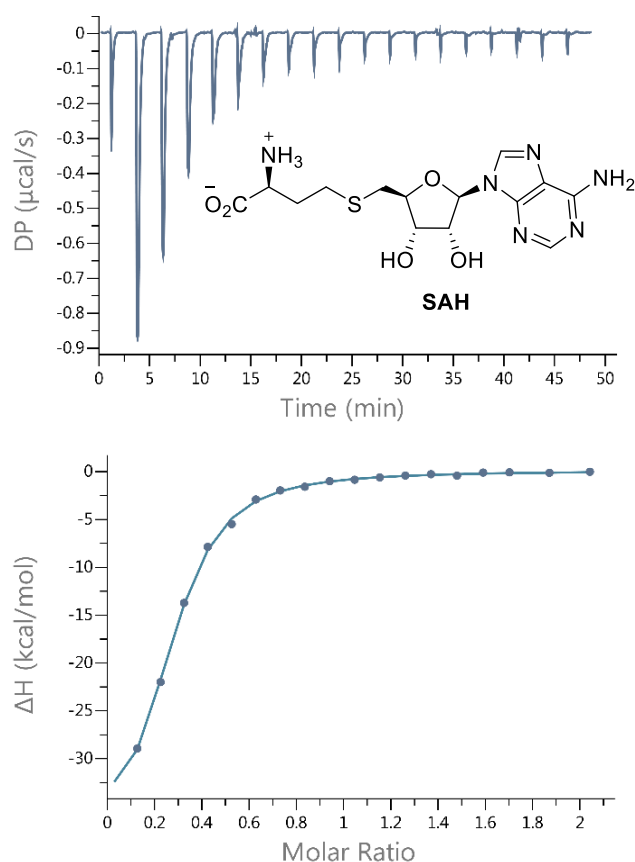


Figure 3-17: ITC results when cofactor SAH ($200 \mu\text{M}$) is titrated into CARM1-2GKG-His ($20 \mu\text{M}$). Raw measured heat changes as a function of time are shown on top, and the integrated measured heats of injection are shown below.

Specified CARM1 + ligand	$K_d \pm$ standard error (μM)	ΔG (kcal/Mol)	ΔH (kcal/Mol)	ΔS (kcal/Mol)
2GKG- CARM1+SAH	1.17 ± 0.12	-7.85	-40.9 ± 1.1	33.1
CARM1+SAH	1.16 ± 0.17	-8.10	-34.6 ± 1.2	11.0

Table 4-6: ITC-derived thermodynamic parameters showed the binding of SAH with human CARM1 when fused with either a cleavable or noncleavable 2GKG soluble tag. The error bars for K_d values are standard errors from curve-fitting.

3.4.2. Crystallisation

Crystallisation trials were performed in order to determine the structure of 2GKG-CARM1 fusion protein. The protein product was incubated with SAH in a 96-well plate at ambient temperature (19 °C) at two different concentrations (7.8 and 4.0 mg/mL) using JCSG+, PEGsII and PACT suite screens. After one month, crystals were obtained from the JCSG+ screen at a protein concentration of 4.0 mg/mL and in conditions of C9 (0.1 M sodium/potassium phosphate, pH 6.5, 25 % (v/v) 1, 2-propanediol), or E10 (1.6 M calcium acetate hydrate, 0.1 M sodium cacodylate 6.5, 40 % (w/v) PEG 300), or D10 (1.6 M magnesium sulfate heptahydrate, 0.1 M MES, 6.5) of the JCSG+screen.

Three crystals were taken to the Diamond Light Source (DLS) for data collection. However, crystal E10 diffracted very weakly, and the others (from condition C9 and D10) gave no X-ray diffraction. Further attempts to co-crystallise the protein in a 24-well plate format with the intention of obtaining better crystals but were not successful in producing any crystals.

Despite not obtaining a solvable crystal structure of CARM1 with the non-cleavable 2GKG soluble tag in complex with SAH ligand, the work documented in this chapter did narrow down possible crystallisation conditions which could be used as templates for further future crystallisation experiments.

3.5. Summary:

The production and purification of CARM1 was optimised with and without cleavable tags at the N-terminus (2GKG tag in pRSF-13 vector) to obtain higher yields of pure protein. Production of CARM1 with a 2GKG soluble tag gave an improved yield when compared with results obtained when using a MBP soluble tag. CARM1 fused with either MBP, or 2GKG soluble tags appeared to give dimeric structures in solution when analysed by gel filtration. The Dimensions obtained from CARM1 with tags, in comparison to crystal structures of homologue domains (PDB code 2Y1X)⁴¹ are consistent with the "ring shape" characteristic of these proteins in solution. The obtained protein was sufficiently stable in solution to allow for crystallisation with bisubstrate inhibitors, which produced useful co-crystal structures with compounds **55**, **58** and **46**. ITC and *in vitro* methylation assays were performed for CARM1, or PRMT1, with H3 and H4 substrate in the presence of bisubstrate inhibitors.

Finally, soluble CARM1 protein was obtained with a non-cleavable 2GKG soluble tag. ITC and crystallisation assays were carried out to determine the binding and the structure of CARM1, with the soluble tag 2GKG. ITC showed strong binding with SAH, which suggest that it folded correctly, but unfortunately, crystals of this co-complex between this protein and SAH were not of sufficient quality to determine their structure.

4. Chapter 4, PRMT6

4.1. Rationale and construct design for PRMT6 (EC: 2.1.1.319)

This chapter focuses on the production and purification of human PRMT6; a protein which was selected to allow for an assessment of the potency and selectivity of PRMT inhibitors, which had previously been tested against PRMT1 and CARM1. PRMT6 has previously been shown to be amenable to a biochemical assay.²⁷ Additionally, attempts were made to produce PRMT6 crystals in complex with our putative inhibitors.

Two constructs were designed and cloned from human PRMT6 cDNA (kindly provided by Dr Clare Davies, University of Birmingham) (Figure 4-2). Regional order neural network (RONN) was used to visualise likely regions of ordered or disordered protein, based on its sequence.¹⁰⁰ This analysis revealed that the first 23 residues are highly glycine-rich and are expected to be disordered. Two constructs of human PRMT6 (residues 28-375 (B) and 39-375 (C)) were selected for cloning and expression; these constructs had their glycine-rich regions deleted with the intention of obtaining stable and folded proteins. Several published crystal structures of PRMT6 also informed this selection, for example, construct (B) (residues 28-375) was chosen through analysis of sequence alignments of the crystal structure derived from *Trypanosoma brucei* PRMT6 (TbPRMT6, PDB code 4LWO).⁴⁶ Glutamate-28 is the residue which precedes the first α -helix in the PRMT6 catalytic domain (Figure 4-1B). This region could introduce unwanted flexibility if the helix is not fully formed. A further shorter construct (C) was designed based on the known structure of the catalytic domain of human (hPRMT6) (PDB code 4Y30)²⁷ the construct which was produced and purified from

insect cells. The crystal structure of hPRMT6 showed that stable, soluble protein could be produced.

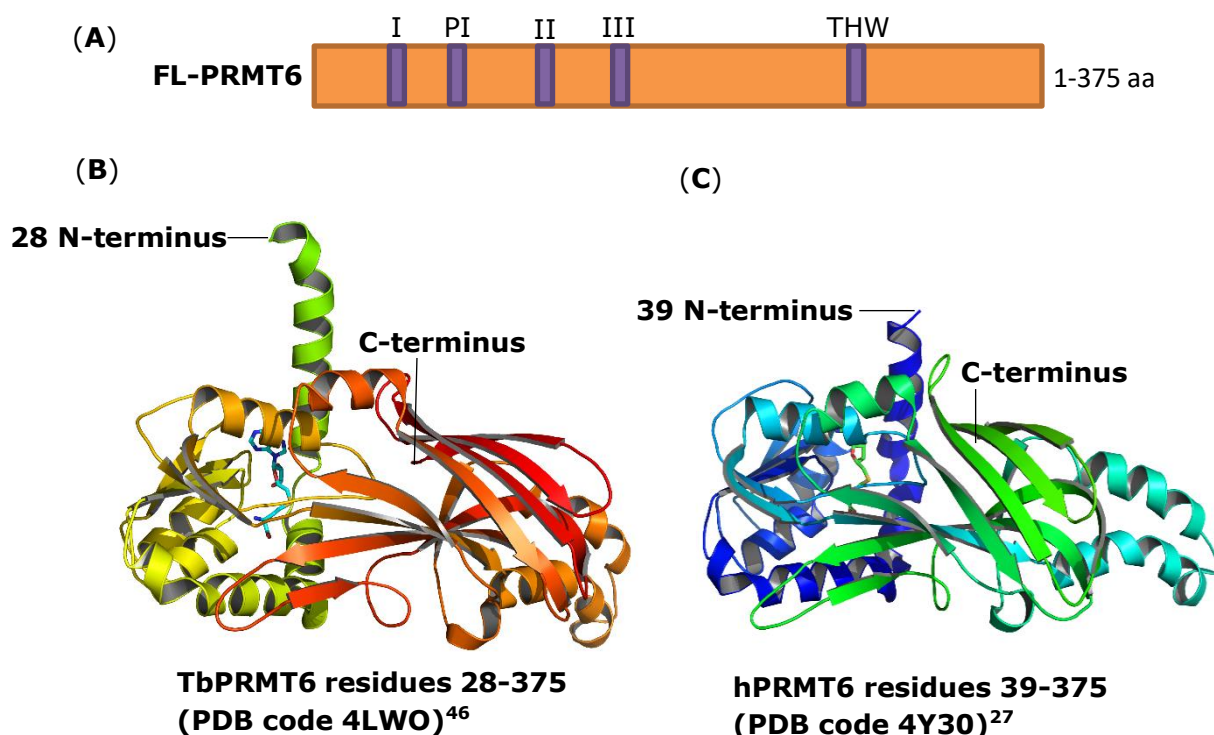


Figure 4-1: PRMT6 catalytic domain structure.

A) Overall domain architecture of the full-length PRMT6. B) Crystal structure of *Trypanosoma brucei* PRMT6 (residues 28-375). C) Crystal structure of human PRMT6 from slightly shorter construct (residues 39-375).

The DNA corresponding to full-length PRMT6 (FL-PRMT6), residues 28-375 (B) and the slightly shorter construct 39-375 (C) were amplified using PCR from cDNA (Figure 4-2). Both PCR products were digested with the appropriate restriction enzymes and ligated into pPROEX HTb vectors.

↓ 28-375 (B)
↓ 39-375 (C)

```

MSQPKRKLES GGGEGGEGTEEDGA ERE AAL ERPRRTKRERDQLY
YECYS DVS VHEEMIADRVRTDAYRLGILRNWAALRGKTVLDVGAGTG
ILSIFCAQAGARRVYAVEASAIWQQAREVVRFNGLED RVHVLPGPVET
VELPEQVDAIVSEWMGYGLLHESMLSSVLHARTKWLKEGGLLLPASA
ELFIAPIS DQMLEWRLGFWWSQVKQHYGVDMSCLEGFATRCLMGHSEI
VVQGLSGEDVLARPQRFAQLEL SRAGLEQELEAGVGGRFRCSCYGS
APMHGFAIWFQVTFPGGESEKPLVLSTSPFHPATHWKQALLYLNEPV
QVEQD TDVSGEITLLPSRDNPRRLRVLLRYKVG DQEEKTKDFAME
DHHHHHH

```

Figure 4-2: PCR amplified the DNA sequence corresponding to full-length PRMT6. Residues 28-375-6×His and a slightly shorter construct 39-375-6×His from the cDNA were chosen for our PRMT6 constructs. The omitted residues are highlighted in red both constructs and blue for the shorter.

4.2. Purification of human PRMT6

4.2.1. PRMT6 construct design

At the beginning of this research, fifteen PRMT6 catalytic domain structures were available in the PDB (e.g., PDB codes 4LWO, 4LWP, 5EGS, 5LV4, 5LV5, 4C03, 4C04, 4Y30). These crystal structures in complex with inhibitors were derived from various organisms and had been produced using different species. This research focused on the domain boundaries 28-375 residues based on the *Trypanosoma brucei* PRMT6 structure (PDB code 4LWO)⁴⁶ and 39-375 residues based on human PRMT6 structure (PDB code 4Y30)²⁷, for our PRMT6 constructs. These constructs were chosen due to TbPRMT6 being successfully produced in *E. coli* and hPRMT6 produced in insect cells, and both displayed *in vitro* methyltransferase activity.^{46,27} Therefore, these constructs were cloned in an attempt to produce soluble PRMT6 protein in *E. coli*.

4.2.1.1. Production and purification of human PRMT6

Two PRMT6 constructs (residues 28-375-6×His and 39-375-6×His) were cloned into pPROEX-HTb which contained a 6×His tag at the N-terminus of the subsequent expression construct (PRMT6-N-His, Figure 4-3). Constructs representing the catalytic domain of hPRMT6 were successfully transformed into LOBSTR BL21 (DE3)-RIL cells.

Due to anticipated difficulties in producing un-tagged PRMT6, a higher concentration of glycerol and salt were utilised in the purification buffers (300 mM NaCl and 5 % (v/v) glycerol), to improve protein stability when compared to another construct. Produced protein construct (residues 39-375) was initially purified using metal chelate chromatography (5 mL HisTrap HP column) and then purified by gel filtration HiLoad 16/16 column (Superdex™ 200 prep grad GE Healthcare). The gel filtration UV trace was characterised by containing at least three overlapping peaks (Figure 4-3). Elution profiles such as these are indicative of the target protein being present at an earlier volume of 45 mL, and therefore containing high concentrations of contaminating proteins. SDS-PAGE analysis revealed aggregated target protein (39.0 kDa) along with a contaminating protein with a molecular weight of 56.0 kDa (Figure 4-3). This construct exhibited poor stability in terms of aggregation.

A possible explanation may be that aggregation was occurring either within the cells as protein was produced, or during purification. Analysis of the protein produced from the other construct containing residues 28-375 gave similar results (data not shown). Two protein constructs analogous to the predicted size (40.0 kDa of protein residues 28-375-6×His and 39.0 kDa of protein residues 39-375-

6×His) were purified, but a poor yield was obtained ~ 0.02 - 0.03 mg/L of bacterial culture.

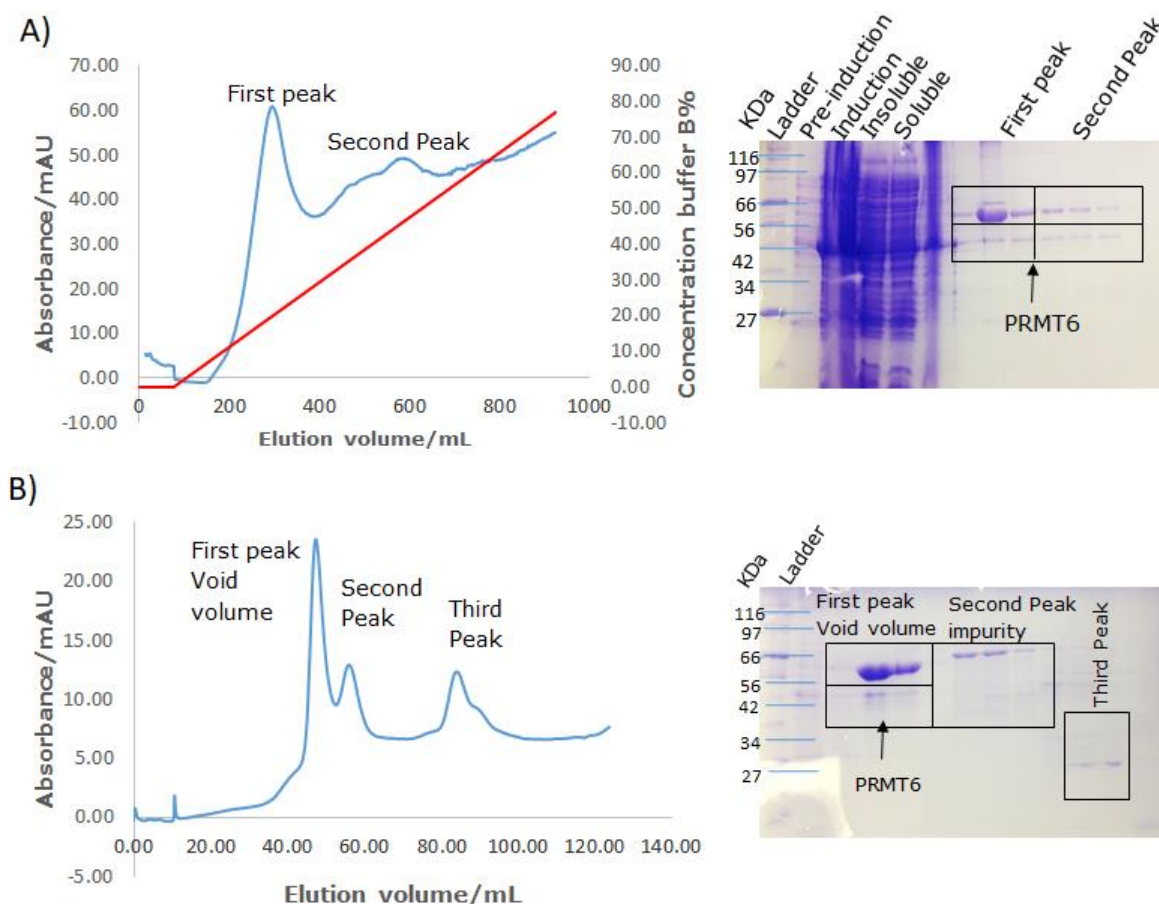


Figure 4-3: Initial purification of PRMT6 (residues 39-375-6×His)

A) UV trace of the nickel affinity chromatography (PRMT6 residues 39-375-6×His in pPROEX HTb vectors) with the peaks from the nickel column are labelled. SDS-PAGE examination of corresponding fractions of the nickel column on the right hand. B) UV trace from a gel filtration with the corresponding SDS-PAGE analysis showed PRMT6 (39.0 kDa) is present as a smaller peak of impurity at 45 mL (Peak 1), and (Peak 2 and 3) contamination was observed as a small peak with the elution volume of 60 and 85 mL respectively.

4.3. Utilising an N-terminus MBP tag optimised for PRMT6 to improve protein purification

An alternative production strategy using a soluble tag was chosen, with a two-pronged approach in terms of improving the solubility of the overall protein construct and increasing the probability of protein crystallisation. Additionally, the choice of the soluble tag would ideally have characteristics that could increase the likelihood of crystallising the fusion protein. Initially, human PRMT6 sequence was cloned into vector pMAL-c2x (E) with the appropriate restriction enzymes, in the same manner as described in chapter three (see Section 3.3.2.). The MBP sequence was modified by Moon *et al.*, to contain five mutations; D82A, K83A, E172A, N173A and K239A (Figure 4-4) which reduced the surface entropy (SER) of the maltose-binding protein (MBP).⁹¹ This approach successfully produced a soluble protein and crystallised several unrelated proteins, which had previously been resistant to crystallisation.¹⁰¹ The MBP sequence was additionally truncated after Asn-367; this truncation removed the TEV cleavage site from the original vector and allowed for the insertion of a short linker region containing three alanine residues, arising from the use of the *NotI* restriction enzyme. The shorter linker region was added to limit flexibility between the MBP tag and the protein.

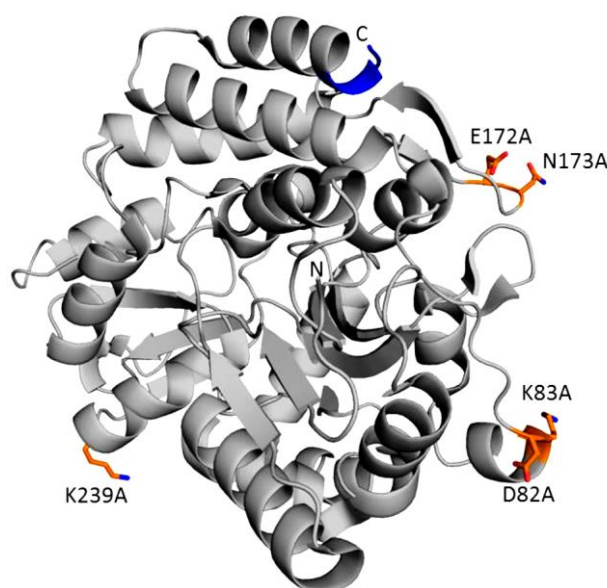


Figure 4-4: Structure of soluble tag maltose-binding protein (PDB code 1HSJ).¹⁰²

Cartoon representation of the structure of the soluble tag maltose-binding protein (MBP) is shown in grey. The C-terminal three alanine residues that form the linker are shown in blue. The five residues are drawn as sticks and labelled (shown in orange) and were subsequently mutated to alanines.

Attempts were made at producing two PRMT6 domain constructs (residues 28-375 and 39-375). These constructs were successfully cloned into the pMAL-c2xE vector with the TEV cleavage site between the MBP fusion and PRMT6 domain boundaries to allow subsequent removal of the MBP tag. As before, a 6×His tag was added to the C-terminus of the PRMT6 sequence to enable the isolation of the fusion protein from the crude lysate by the nickel affinity chromatography.

Test production of these constructs was performed in an attempt to obtain an active sample of MBP-PRMT6×His. These constructs were produced in the same manner as before and then purified using nickel affinity chromatography and the gel filtration column (GE Healthcare). Both constructs containing a MBP tag were produced in excellent initial yield. Unfortunately, aggregation was again observed

for both constructs based on the elution volume arising from the gel filtration (data not shown).

Digest with TEV protease to ensure MBP tag cleavage from the target protein was also examined. The proteins were successfully cleaved (Figure 4-4). A second gel filtration was performed to separate PRMT6-His from the MBP tag and TEV protease. However, the PRMT6 construct (39.0 kDa, residues 39-375-6×His) aggregated and eluted in the void volume at 45 mL (Figure 4-4). The MBP tag eluted at 82 mL corresponding to the monomeric MBP tag (41.0 kDa)⁹⁰ (as judged by gel filtration and SDS PAGE, Figure 4-4). Analysis of the produced protein from the other construct (residues 28-375-6×His) was not shown as results were similar. The yield obtained for both proteins was approximately 0.17-0.2 mg/L.

Another possible purification of protein was attempted by adding 1 % (v/v) TWEEN® to the buffer solution. It has been suggesting that TWEEN® could help prevent aggregation of the protein during purification by increasing hydrophobicity of the solution.⁹³ Unfortunately, no improvement was observed in this case.

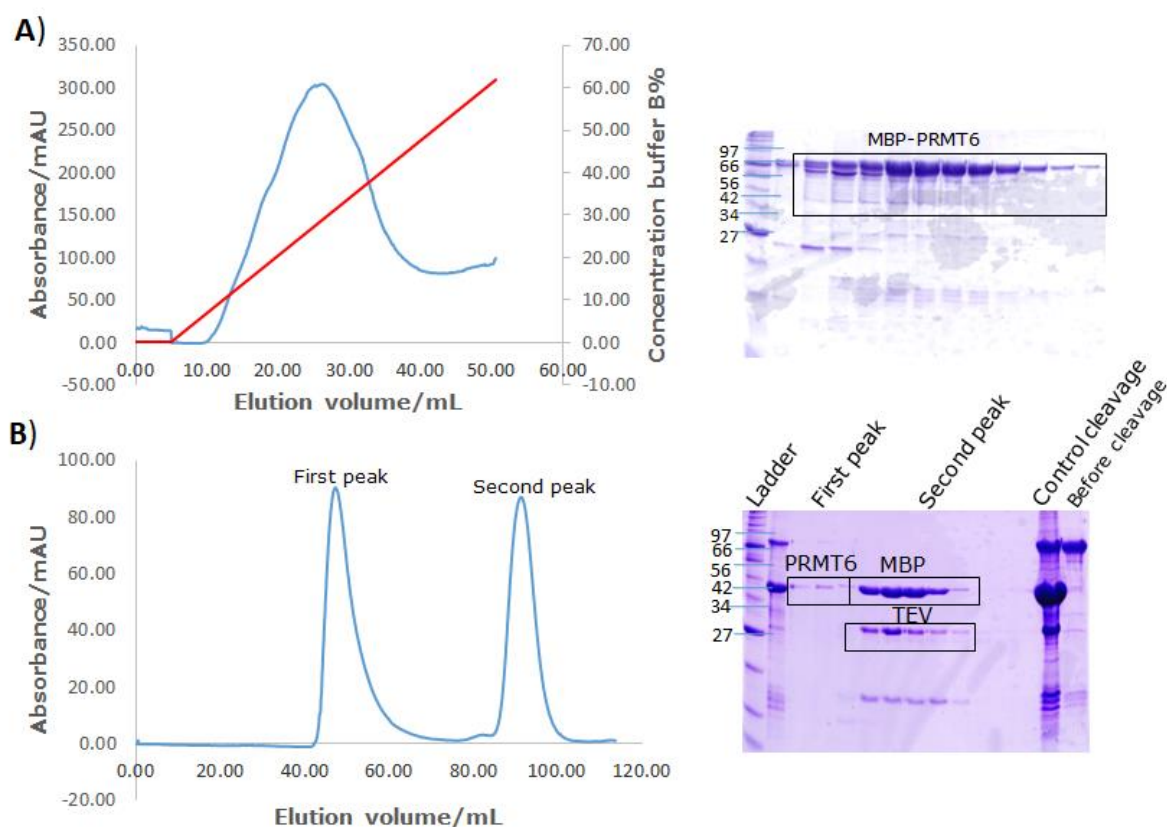


Figure 4-5: Purification of MBP-PRMT6-His (residues 39-375-6×His) constructs
 A) UV trace of the nickel column of MBP-PRMT6-His with the subsequent SDS-PAGE analysis of the fractions from chromatography on the right. B) UV trace of gel filtration of MBP with the corresponding SDS-PAGE. PRMT6 after cleavage obtained an aggregation product at approximately 45 mL, while MBP (41.0 kDa) appears as a monomer product at about 82 mL.

4.3.1. PRMT6 catalytic domain-soluble tag fusions

A further strategy was investigated using a different small soluble tag when expressing PRMT6 with the aim of producing a soluble protein. The 2GKG tag is a 14.0 kDa receiver domain derived from *Myxococcus xanthus* social motility protein frzS (PDB code: 2GKG, Figure 4-6)⁹³ and the 1BKR tag is a 13.3 kDa calponin homology domain from human β -spectrin (referred to as the PDB code: 1BKR, Figure 4-6).¹⁰³ These tags were previously identified through analysis of the crystal

structure database⁹⁴ and have been successfully used for the purification of a different protein (Dreveny group, unpublished results). These tags are less than 400 amino acids in size and do not contain disulfide bonds. 1BKR tag and 2GKG can be produced well in *E. coli*.^{93,103}

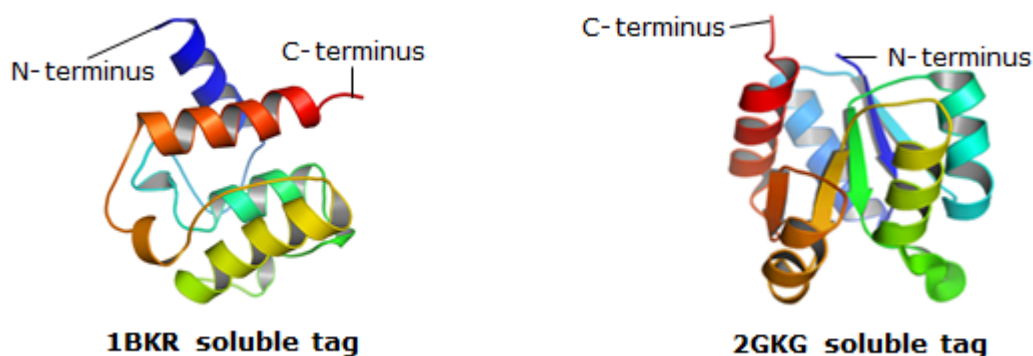


Figure 4-6: Cartoon structures of 1BKR¹⁰³ and 2GKG.⁹³

4.3.2. Test production of PRMT6 with fusion tag 2GKG and 1BKR

A 2GKG fusion tag had been used previously for the production of CARM1 and gave encouraging results (see Section 3.3.4, Chapter Three). The 2GKG fusion protein was therefore chosen to produce PRMT6 protein by inserting it at the N-terminus, with the TEV cleavage site, to facilitate later removal of the 2GKG tag. Unfortunately, when using this tag with the PRMT6 expression clone, the produced protein eluted in the void volume during gel filtration, due to aggregation (data not shown). This caused significant issues for production of the protein, especially for protein activity and crystallisation.

Further work was conducted toward developing a reliable method for generating a soluble PRMT6 protein. Attention turned to another soluble tag 1BKR¹⁰³ which

has been successfully used to improve solubility and protein folding (as described in the following sections 4.3.3).

4.3.3. Utilising an N-terminus 1BKR tag optimised for hPRMT6

This experiment sought to produce and purify PRMT6 as a 1BKR fusion protein to allow for the correct folding of the two domains. 1BKR was chosen because the presence of this tag has been shown to improve protein solubility when used as a fusion partner.⁹⁴ This tag, when fused to several different types of protein, has been shown to produce a soluble protein in several cases (Dreveny group, unpublished results).

To explore this possibility, the production and purification of PRMT6 construct (residues 28-375-6×His) with this tag was optimised. Also assessed was whether the presence of the tag could aid in the proper folding of recombinant proteins. PRMT6 was connected to 1BKR at the N-terminus, with the TEV cleavage site, using the *NotI* restriction enzyme between two units during cloning. This cloning strategy resulted in three alanine residues present as a short linker sequence between PRMT6 and 1BKR tags (Figure 4-7). Initially, the presence of these three residues was thought to be sufficiently flexible to enable the folding of the target proteins. The catalytic domain PRMT6 sequence (residues 28-375-6×His) was, therefore, sub-cloned into the pRSF-13 vector with the TEV cleavage site between two units.



Figure 4-7: 1BKR fusion connected to PRMT6 constructs 28-375-6×His.

Human PRMT6 with the cleavable 1BKR soluble tag in the pRSF-13 vector was transformed into LOBSTR BL21 (DE3)-RIL, cultures grown, and the protein expressed. Soluble recombinant protein obtained from the subsequent protein purification was found to have successfully improved solubility and had undergone proper folding. Analysis by gel filtration (S200 gel filtration column) showed that oligomers 1BKR-PRMT6 (56.0 kDa) were obtained from a peak with an elution volume of 62 mL (Figure 4-8). Aside from PRMT1, there is no published data which suggests that PRMT6 is present as an oligomer; the native form of this protein appears to be a dimer.¹⁰⁴ Analysing the gel filtration elution profile meant it was difficult to distinguish between high molecular weight oligomers and PRMT6 aggregates based solely on their gel filtration elution profiles. Additionally, a preliminary *in vitro* methylation assay was carried out by Dr Emma Gunnell to test the activity of this purification, but no methyltransferase activity was detected. The total yield obtained from the subsequent protein purification was 34 µg/L. Therefore, the tag-free protein was generated using a tobacco etch virus (TEV) protease to remove the fusion 1BKR tag from target proteins, which are discussed below.

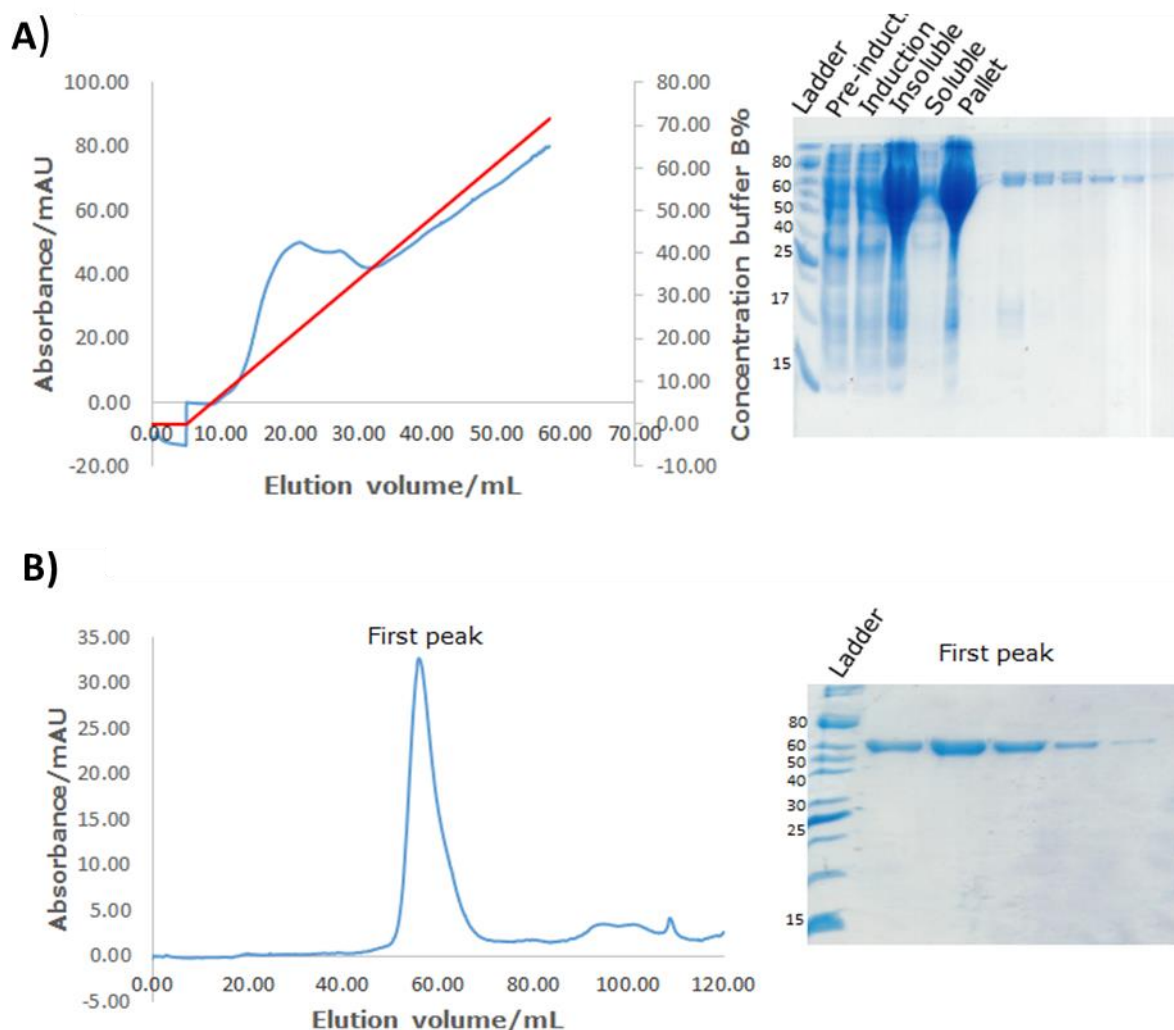


Figure 4-8: purification of 1BKR-PRMT6-His.

A) UV trace of the nickel column of 1BKR-PRMT6-His with the corresponding SDS-PAGE gel analysis. Fractions from 20–40 mL were combined and concentrated for subsequent gel filtration. B) The gel filtration (Superdex 200 16/60) chromatogram with the corresponding SDS-PAGE of fractions eluting from the gel filtration. The major species were eluted from the gel filtration as oligomer has a molecular weight of ~56.0 kDa.

A small-scale optimisation was carried out using 0.01 to 0.1 molar equivalents of TEV protease to cleave the soluble tag from PRMT6. This was performed to determine how much the TEV enzyme was needed to cleave the fusion protein. Preliminary analysis of the attempted proteolytic cleavage of the soluble tag showed that approximately 90 % of the protein was not cleaved (Figure 4-9). It

may be that TEV protease did not recognise the cleavage site in this fusion protein. Another possible explanation for this is that some intermolecular interaction between residues of 1BKR and the catalytic domain is facilitated due to the flexible nature of the linker, which may serve to protect the site from cleavage. However, further work would need to be undertaken to confirm this.

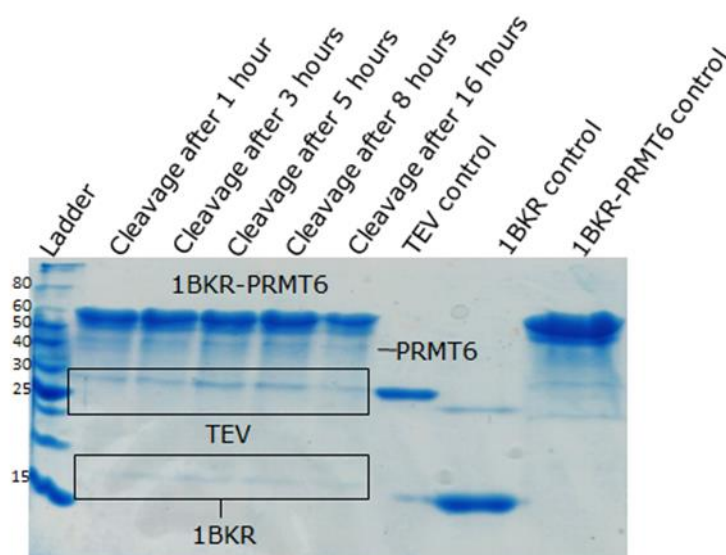


Figure 4-9: Time course analysis for 1BKR-PRMT6-His.

SDS-PAGE showing different times and a different ratio of TEV protease to cleave the 1BKR tag from PRMT6-His at room temperature.

As described on the previous page, it is possible that the TEV cleavage site facilitates obstruction of PRMT6 active site and may influence both the model ligand (SAM) or substrate binding to the active site of the protein. It was decided to remove the TEV cleavage site and attempt production of a construct without a TEV cleavage site between PRMT6 (residues 28-375-6×His) and the 1BKR soluble tag. The construct was designed with the aim of limiting the flexibility between the 1BKR tag and protein.

Additionally, it was also previously observed that 1BKR-PRMT6-His was produced in *E. coli* LOBSTR BL21 (DE3)-RIL cells at a low yield (Figure 4-8) To solve this problem, Arctic Express cells which produce an additional chaperonin were utilised to increase the yield and the solubility of the produced protein.¹⁰⁵ To explore this possibility, a PRMT6 with a non-cleavable tag construct (residues 28-375-6×His) was produced and purified to maximise the likelihood of later obtaining a higher yield of the stable and folded protein for use in inhibitor binding studies.

4.4. 1BKR-PRMT6-His expression in Arctic Express cells

A practical solution to the problem of protein aggregation and low yield is to co-produce with a chaperonin in *E. coli*.¹⁰⁵ Arctic Express competent cells contain the chaperonin system Cpn60/10 from the psychrophilic bacterium *Oleispira antarctica*; the presence of this chaperonin means that the system remains active at low temperatures.^{105,106} Production of protein by using chaperonins is often carried out at low temperatures (12–15 °C) to improve the solubility of the proteins and allowing time (~72 hrs) for adequate folding protein and improved stability.¹⁰⁵

Co-production of PRMT6 with the non-cleavable 1BKR soluble tag (residues 28-375-6×His) in *E. coli* strain containing Cpn60/10 chaperonin was successful, leading to initial yields of 0.47 mg/L of bacterial culture, with soluble protein being present as an oligomer (Figure 4-10). Initial purification suggested that the target 1BKR-PRMT6-His co-purified with the Cpn60/10 chaperonin. This proposed assignment was confirmed using mass spectrometry, which was discussed in Section 4.4.2. Figure 4-12. Western blot was also carried out to confirm the

presence of the histidine tag of our target protein (Figure 4-11). Isothermal titration calorimetry (ITC) analysis of the binding of SAH to this protein was also performed and will be discussed in more detail in Section 4.4.5. (Figure 4-15).

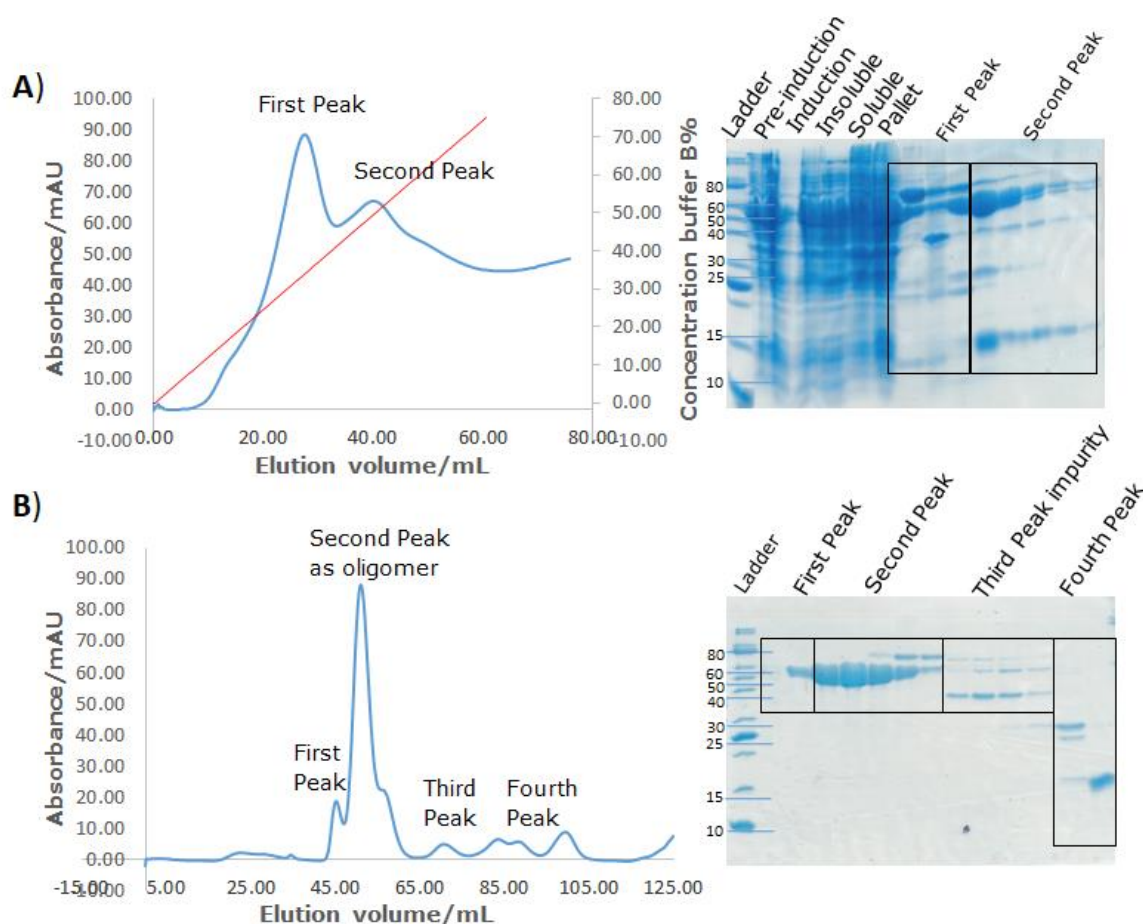


Figure 4-10: Purification of co-production 1BKR-PRMT6-His with chaperonin.

A) UV trace of the nickel affinity chromatography of 1BKR-PRMT6-His with the corresponding SDS-PAGE analysis of the fractions, samples are taken from pre-induction, induction, soluble protein, unbound protein, and insoluble protein, respectively. Fractions from second peaks were combined and concentrated for subsequent gel filtration. B) UV trace of gel filtration of 1BKR-PRMT6-His with chaperonin shows two major peaks which indicated PRMT6 eluted as an aggregation protein and a second peak as an oligomer. SDS-PAGE analysis for the gel filtration was taken from the concentrated fractions of the purification step.

4.4.1. Western blot used for 1BKR-PRMT6-His

A western blot was performed to examine purified protein samples using positive and negative control samples for comparison (Figure 4-11). PRMT1 containing a 6×His tag was used as a positive control, and GSTUCL1 was used as a negative control, as this construct does not include a 6×His-tag (GSTUCL1 from PhD student Camila Bonturi). This analysis confirmed the presence of the 6×His tag in the isolated protein and a protein band was observed which corresponded to the native *E. coli* 1BKR-PRMT6-His. This confirmed that the construct was present in the clarified lysate.

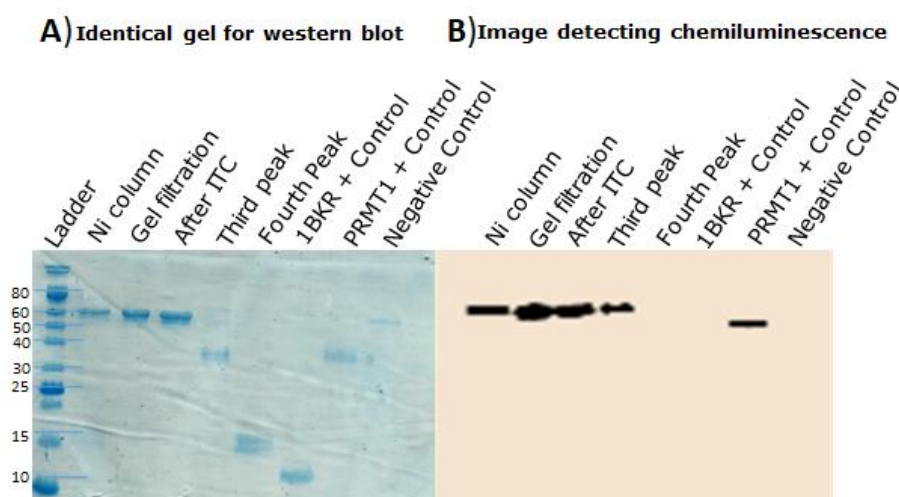


Figure 4-11: Western blots confirmed that the purified protein corresponded to 1BKR-PRMT6-His. A) SDS page of PRMT6 with 1BKR samples after purification on the nickel column and gel filtration, example after running ITC, 1BKR positive control, PRMT1 positive control and negative control (GSTUCL1 from PhD student Camila Bonturi). B) Western blot of fractions eluting from purification column, using primary antibody (primary mouse anti-tetra-His antibody) and secondary antibody (Secondary anti-mouse horseradish peroxidase (HRP) conjugated antibody).

4.4.2. Mass spectrometry

Purified protein was analysed by SDS-PAGE and the protein band expected to contain our target protein was excised from the gel. The excised band was sent to mass spectrometry analysis at the University of York using a MALDI-TOF,^{107,108} to identify the composition of the protein. The results of mass spectrometry analysis show that fragments with masses highlighted (Figure 4-12) were identified and then subject to a BLAST search using the UniProt online software to confirm the identity of the parent peptide. The overall results demonstrated that each peptide corresponded to either PRMT6 or the chaperonin Cpn60/10.¹⁰⁹ The result indicated that two peptides (orange in Figure 4-12) correlated with PRMT6 and three peptides (red) correlated with chaperonin Cpn60/10.

Sequence of protein chaperonin Cpn60/10.

MAAKDVLFGD SARA**KMLVGV NILADAVRV**T LGPKGRNVVI EKSFGAPIIT
KDGVSVAIEI ELKDKFENMG AQMVKEVASQ ANDQAGDGT TATVLAQAI
SEGLKSVAAG MNPMDLKRGI DKATAAVVAA IKEQAQPCLD TKAIQVGTI
SANADETVGR LIAEAMEKVG KEGVITVEEG KGLEDELVDV EGMQFDRGYL
SPYFINNQEK MTVEMENPLI LLVDKKIDNL QELLPILENV **AKSGRPLLI**V
AEDVEGQALA TLVVNNLRGT FKVAAVKAPG FGDRRKAMLQ DLAILTGGQV
ISEELGMSLE TADPSSLGTA SKVVIDKENT VIVDGAGTEA SVNTRVDQIR
AEIESSTSDY DIEKLQERVA KLAGGVAVIK VGAGSEMEMK EKKDRVDDAL
HAT**RAAVEEG VVAGGGVALI RALSSVTVG DNEDQNVGIA LALRA**MEAPI
RQIAGNAGAE GSVVVDKVKV GTGSFGFNAS TGEYGDMIAM GILDPKAVTR
SSLQAAASIA GLMITTEAMV ADAPVEEGAG GMPDMGGMGG MGGMPGMM

Sequence of protein arginine methyltransferases 6

MSQPKKRKLE SGGGGEGGEG TEEEDGAERE AALERPRRTK RERDQLYYEC
YSDVSVHEEM IADRVRTDAY RLGILRNWAA **LRGKTVLDVG AGTGILSIFC**
AQAGARRVYA VEASAIWQQA REVVRFNGLE DRVHVLPGPV ETVELPEQVD
AIVSEWMGYG LLHESMLSSV LHARTKWLKE GLLLLPASAE LFIAPISDQM
LEWRLGFWSQ VKQHYGVDMV CLEGFATRCL MGHSEIVVQGLSGEDVLARP
QRFAQLELSR AGLEQELEAG VGGRFRCSCY GSAPMHGFAI WFQVTFPGGE
SEKPLVLSTS PFHPATHWKQ ALLYLNEPVQ VEQDQDVSSE ITLLPSRDNP
RRLRVLLRYK VGDQEEKTKD FAMED

Figure 4-12: Mass spectrometry data identifying the composition of the protein.

The above sequence highlighted in the red was matched to the protein sequence 60.0 kDa chaperonin (OS=Oleispira antarctica GN=groL). The bottom sequence highlighted in the orange was matched to a sequence of PRMT6.

Removing the chaperonin Cpn60/10 from the mixture with PRMT6 was necessary due to the possible effect on the activity of the protein. Further work was therefore required to remove the co-purified chaperonin from 1BKR-PRMT6-His due to specific binding interaction. The co-purified protein was isolated and removed from the fusion protein, as discussed below.

4.4.3. Removal of co-purifying chaperonin Cpn 60/10¹⁰⁹

Co-purification of chaperonins is a common problem that can limit the activity of the protein after production,¹⁰⁵ incubation with a buffer containing (5 mM MgCl₂, 5 mM ATP, 300 mM NaCl and 20 mM HEPES-NaOH (pH 7.0)) has been shown to allow a change of chaperonin conformation and release from the target protein.¹⁰⁵ Incubation of co-purified chaperonin with 1BKR-PRMT6-His using the above mixture effectively removed the Cpn60/10 chaperonin from PRMT6 immobilised on the nickel affinity chromatography. As shown in Figure 4-13, chaperonin eluted from 1BKR-PRMT6 using the wash buffer with a molecular weight of approximately 60.0 kDa as observed by SDS-PAGE, while a band which had eluted from the nickel column with molecular weight 56.0 kDa corresponding to 1BKR-PRMT6-His was also produced. The fraction from the nickel column was concentrated and then subjected to gel filtration (Superdex 200 16/60) to isolate any impurities or aggregated protein. Analytical gel filtration showed that the first eluting peak consists of a large amount of aggregated protein, whereas the later peak contained 1BKR-PRMT6-His protein eluted at approximately 82 mL, as a monomer. To ensure that the bacterially produced 1BKR-PRMT6-His domain was active as a monomer, a well-established *in vitro* methylation assay was used to test catalytic function. The result of this investigation confirmed that PRMT6 catalytic domain

was a monomer and was active for methylation of target substrate histone peptide H3R3. Activity assay of this purification is discussed in Section 4.4.7.

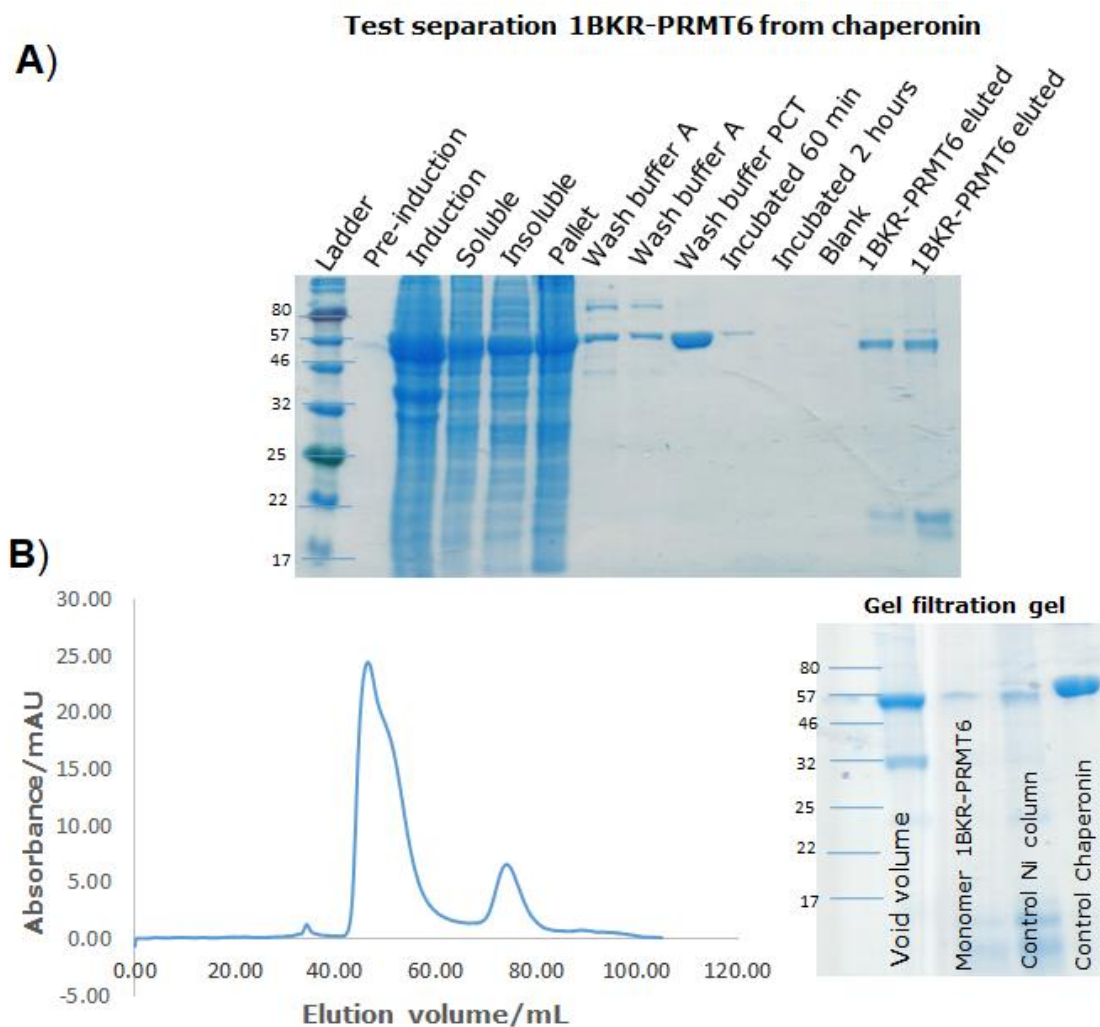


Figure 4-13: Removal of co-purifying chaperonin from 1BKR-PRMT6-His.

A) SDS-PAGE analysis showing co-purification of chaperonin impurity separated from 1BKR-PRMT6-His by incubated with a buffer containing (5 mM MgCl_2 , 5 mM ATP, 300 mM NaCl and 20 mM HEPES-NaOH (pH 7.0)). B) UV trace of gel filtration for 1BKR-PRMT6-His and the corresponding SDS-PAGE analysis. The first band on the SDS-PAGE gel corresponded to the first peak of gel filtration column and eluted as an aggregated protein. The second band represents the elution peak as a monomer. The last two bands represent the positive control of both separation chaperonin and sample of 1BKR-PRMT6-His from the nickel column.

The product of this construct (1BKR-PRMT6-His residues 28-375-6×His) was encouraging as the protein was soluble and active, without the TEV cleavage site. This suggests that the removal of the TEV cleavage site that joins the two domains could be necessary as this is a flexible region.

However, the yield obtained from the purification of PRMT6 as a monomer was 0.042 mg/L of bacterial culture. This yield did not give sufficient protein to carry out binding studies using ITC, as this requires at least 1 mg/L of 1BKR-PRMT6-His to obtain a curve with sufficient sigmoidicity for an accurate K_d determination. Consequently, autoinduction media was chosen,^{105,106} as protein production at low temperatures can improve protein yield and further enhance the stability and solubility of the protein.¹⁰⁶ This strategy had previously increased the yield obtained from the production of various related proteins.¹⁰⁶

4.4.4. Autoinduction at a higher cell density increases protein yield¹⁰⁶

Autoinduction media was used to produce high densities of the cells in *E. coli* in order to obtain a higher yield of the protein. The protein-coding sequence was cloned into the plasmid under control of a T7 lac promoter. Expression from a single colony was grown in non-inducing media, which meant that no IPTG induction was required. Auto-induction can be regulated by adjusting glucose/lactose levels in media at the first stage of cell growth. (Further described in chapter 7).^{105,106}

Non-cleavable IBKR-hPRMT6 (residues 28-375-6×His) was transferred to Arctic Express competent cells, and protein production carried out using autoinduction

media and purified as before. Analysis of the purified fractions revealed a large amount of aggregated protein in the first peak, but a suitable amount of protein eluted as a monomer in the second fraction. The yield of protein produced from autoinduction media was approximately 0.18 mg/L of bacterial culture, which compared favourably with the yield obtained from LB media (0.042 mg/L). A possible explanation as to why autoinduction media appeared to have improved the yield of protein, could be that induction gradually increased in the autoinduction media rather than rapidly upon addition of IPTG during expression.¹⁰⁶ Examination of fractions obtained from the nickel and the gel filtration column by SDS-PAGE suggested that 1BKR-PRMT6-His was ~80 % pure. (Figure 4-17).

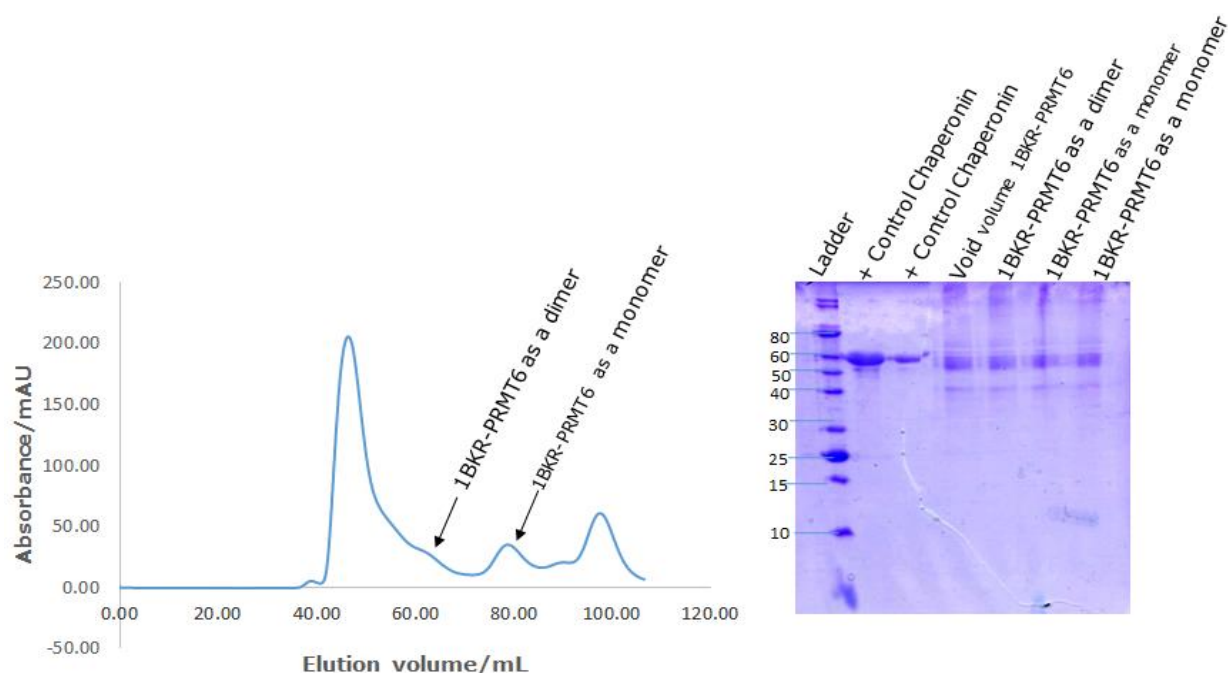


Figure 4-14: Purification of 1BKR-PRMT6-His using autoinduction media.

UV trace of gel filtration for 1BKR-PRMT6-His with the corresponding to SDS-PAGE analysis. The gel showed the band from each peak of gel filtration with chaperonin as a control taken from previous purification step.

4.4.5. Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC)¹¹⁰ was performed to determine the binding affinity of SAH with the purified 1BKR-PRMT6-His. Two separate experiments were performed. The initial investigation was carried out for the first purification of 1BKR-PRMT6-His, which had co-purified with chaperonin (Section 4.4, Figure 4-10). A second experiment was performed with the fusion protein with chaperonin removed as discussed above. The K_d of SAH for these purified proteins was determined by curve fitting using a binding model in the MicroCal PEAQ-ITC Analysis Software.

An initial titration was performed in which aliquots of SAH ligand (290 μ M) in the reaction buffer containing (150 mM NaCl, 1 % (v/v) glycerol, 20 mM Tris-HCl and pH 8.0) was injected into a 29 μ M solution of 1BKR-PRMT6-His containing chaperonin present in the mixture, alongside a control titration with the corresponding ligand-to-buffer. Both ligand and target protein were dialysed into the same buffer (150 mM NaCl, 1 % (v/v) glycerol, 20 mM Tris-HCl and pH 8.0). This work showed that 1BKR-PRMT6 formed a relatively weak interaction with the ligand (SAH, K_d 81 ± 12.0 μ M) when compared with published data for SAH-TbPRMT6 which had a K_d of 45 ± 5.0 μ M⁴⁶ (Figure 4-15). A second similar titration was performed with chaperonin removed to investigate effects binding of SAH into the 1BKR-PRMT6 as a monomer protein. Intriguingly, this 'purified' 1BKR-PRMT6-His was found to bind strongly to SAH with a K_d of 3.30 ± 0.9 μ M. Comparing the two results, it can be seen that there was an increased affinity of SAH for 1BKR-PRMT6-His as a monomer which bound strongly when compared to data from the literature, SAH-TbPRMT6 has a K_d of 45 ± 5.0 μ M⁴⁶ and purified protein with chaperonin has a K_d of 81 ± 11 μ M. The increased affinity of SAH into 1BKR-

PRMT6-His as a monomer versus chaperonin purified may be due to the chaperonin obstructing PRMT6 binding site; this would be interesting to explore further in the future.

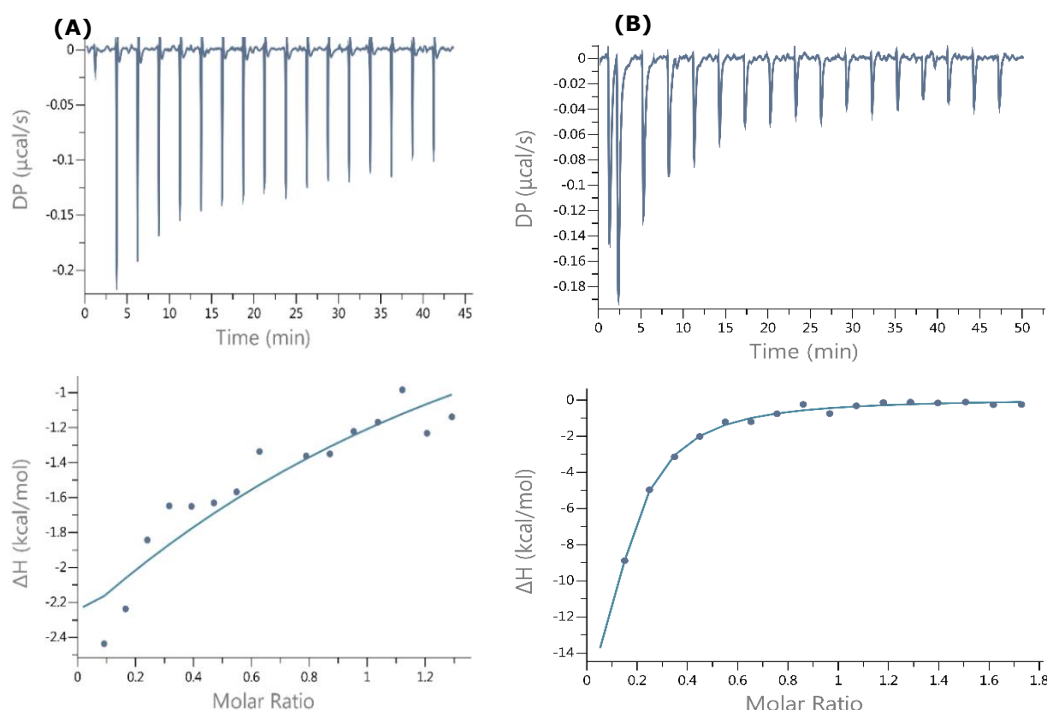


Figure 4-15: Isothermal titration calorimetry (ITC) data for the titration of SAH into 1BKR-PRMT6-His.

A) Experiment injecting 8 μL aliquots of a solution of 290 μM SAH into a 29 μM solution of 1BKR-PRMT6-His with co-purified chaperonin. The top isotherm panels A and B are the raw ITC data showing the exothermic release of heat upon SAH binding, and the bottom panels show the isotherm produced when integrating the peaks. (B) ITC data showed binding of SAH with 1BKR-PRMT6-His as a monomer in solution.

Protein + ligand	Kd $\pm$ standard error (μM)	ΔG (kcal/Mol)	ΔH (kcal/Mol)	ΔS (kcal/Mol)
PRMT6 with chaperonin + SAH	81.3 $\pm$ 11.8	-5.58	-69.2 $\pm$ 58.8	1.38
PRMT6+ SAH	3.30 $\pm$ 0.9	-7.48	-42.4 $\pm$ 30.7	34.9

Table 4-1: ITC data for two purifications of 1BKR-PRMT6-His.

ITC-derived thermodynamic parameters showed the data binding when titrating the SAH into each PRMT6 with chaperonin and PRMT6 as a monomer. The error bars for Kd values are standard errors derived from curve-fitting.

4.4.6. Evaluation of human 1BKR-PRMT6 catalytic activity

Upon obtaining pure and monomeric 1BKR-PRMT6-His (see Section 4.4.4.), it was of interest to determine if this protein was catalytically active. A radiometric activity assay was performed to assess the enzymatic activity of 1BKR-PRMT6-His. Histone-3 (residues 1-21, as GB-fused peptide) was produced and purified as described below (Figure 4-16) for use as substrates, to test the activity of a purified protein. This was because PRMT6 had been shown to methylate arginine two on histone 3 (H3R2) *in vivo*.^{111,17} Brief optimisation of this assay is detailed in the following section.



Figure 3-16: Sequence representation of the peptides from histone 3.

The blue squares represent the enzymes of PRMT6 methylate the arginine residues highlighted in light blue which lie within the H3 boundaries. The amino acid sequences for this peptide (residues 1-22 H3) are cloned into the expression vector.

Due to the small size of the non-tagged histone 3 peptide (residues 1-21, 4.0 kDa), this peptide was linked to GB fusion tag (6.0 kDa) to allow the production of peptides on the nickel column and aid to visualisation using SDS-PAGE. Histone 3 peptide (residues 1-21) was successfully cloned as a GB-pET21a fusion construct and transformed into BL21-CodonPlus (DE3)-RIL competent cells (Agilent). Initial production of GB-fused histone 3 peptide (6×His-GB-H3 (residues 1-22)) was carried out at 37°C in 1L of LB media. Protein was then purified in the same manner as previously. Initial yields obtained of the peptide were excellent, 3.3 mg/L of bacterial culture and soluble protein was obtained (Figure 4-17).

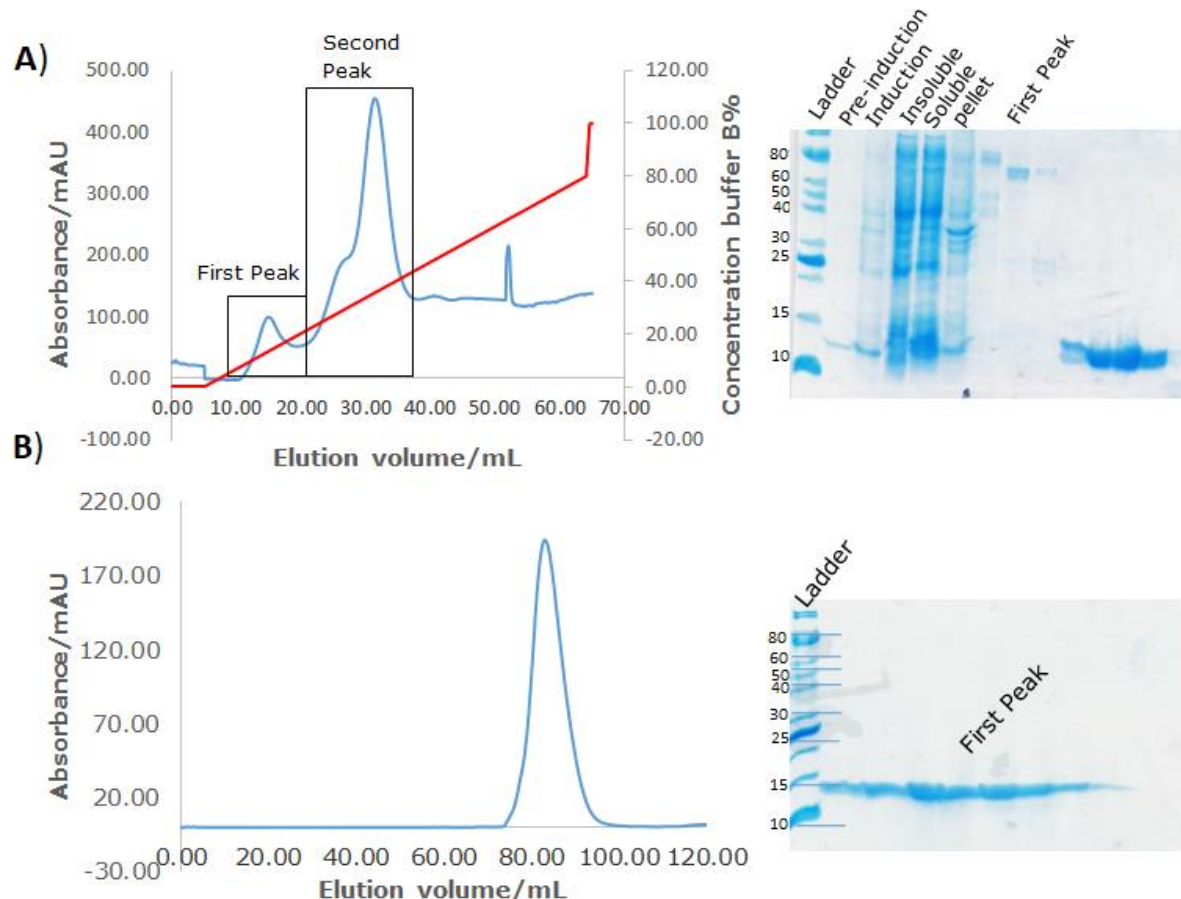


Figure 4-17: Purification of peptide histone 3 GB-H3 (residues 1-22)

A) UV trace of the nickel affinity chromatography eluent representing peptide histone 3 GB-H3 (residues 1-22) with the corresponding SDS-PAGE analysis of the fractions. B) UV trace of gel filtration with the relevant protein bands on the gels filtration is labelled.

4.4.7. Radiometric (^3H) activity assays

Radiometric activity assays were performed to determine if purified non-cleavable 1BKR-PRMT6-His construct was catalytically active.⁴⁴ Experiments conducted by Dr Emma Gunnell evaluated monomeric 1BKR-PRMT6-His (16 μM) with the substrate concentration tail region (163 μM) in the presence of [^3H]-SAM (0.23 μM). A histone concentration of 163 μM was initially chosen arbitrarily due to PRMT6 apparent automethylation capability.³⁴ Initial results revealed that 1BKR-PRMT6-His possessed methyltransferase activity with substrate H3, which was

comparable to PRMT1 with H4 substrate used as a positive control (Figure 4-18). The omission of substrate H3 from the assay showed that, unlike PRMT1, IBKR-PRMT6 is capable of auto-methylation (Figure 4-18), which is consistent with published findings.³⁴

The second test examined 1BKR-PRMT6-His at 17.0 μ M as a dimer, which was observed on the gel filtration at 62 mL (see Section 4.4.4, Figure 4-14). Unexpectedly, the dimer methyltransferase activity was not observed. These results may be due to the presence of the chaperonin impurity still binding to the dimer of 1BKR-PRMT6-His, or interference from the 1BKR tag. Our findings were plotted as disintegrations per minute¹ (DPM¹) against time with the response rate of 1BKR-PRMT6-His activity occurring within the initial 60 minutes.

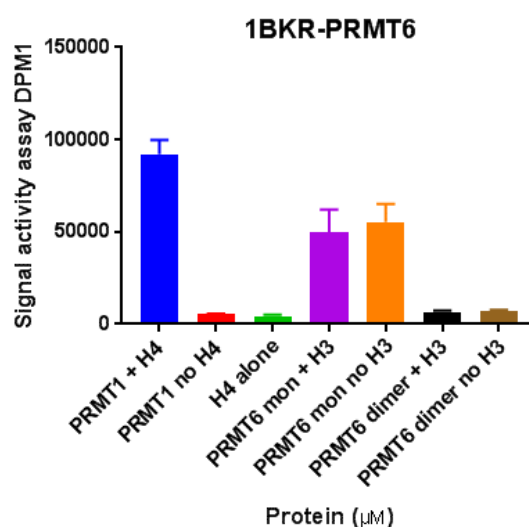


Figure 4-18: Initial activity test of 1BKR-PRMT6-His.

DPM indicates the disintegrations per minute recorded by liquid scintillation counting and is proportional to histone methylation. Results from *in vitro* methylation assay, PRMT1 with the substrate H4 as a positive control (column 1), one in the absence of H4 substrate as a negative control (column 2) and lacking PRMT1 as a negative control (column 3). 1BKR-PRMT6-His was incubated with ³H-SAM in the presence of H3 (column 4) and in the absence of H3 (column 5) demonstrate its automethylation capability. Dimeric 1BKR-PRMT6-His in presence and absence of H3 (column 6 & 7 respectively). Error bars indicate the standard deviation from 3 repeat determinations.

4.4.8. Linear range determination for 1BKR-PRMT6-His

Further, methylation assays were carried out by Dr Emma Gunnell to determine the range of concentrations for signal varied linearly with 1BKR-PRMT6-His (Figure 4-19). 1BKR-PRMT6-His protein concentration used in these experiments was optimised by initially testing concentration ranges from 23 nM -2.5 μ M, alongside constant substrate concentrations of 76 μ M in the presence of [^{3} H]-SAM 0.23 μ M. The assay signal was found to vary linearly up to > 1.5 μ M 1BKR-PRMT. This concentration of 1.5 to 2 μ M was chosen in further assays with bisubstrate inhibitors. The data were analysed using GraphPad Prism with one site-specific binding (Figure 4-19).

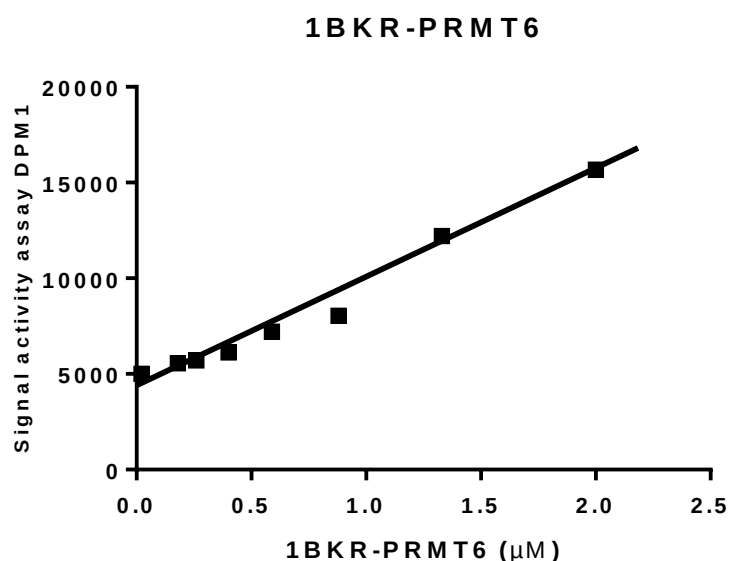


Figure 4-19: Determining the linear range of 1BKR-PRMT6-His assay signal. No error bars are shown as this prelim experimental was only performed once.

4.5. Test inhibition of small molecule inhibitors against hPRMT6

Having established that inhibitors **55**, **58**, **46**, and **64** were all selective for CARM1 over PRMT1 as discussed in chapter three (see Section 3.2.2. and 3.3.5) how these

inhibitors bound to PRMT6 was investigated. Preliminary experiment on the inhibition of PRMT6 was conducted by Dr Emma Gunnell. To do this, an *in vitro* methylation activity assay was performed to evaluate the effect of inhibitors on IBRK-PRMT6-His protein. Bisubstrate inhibitors **55**, **58**, **46**, **64** and **39** concentrations ranging from 1 mM, 100 μ M and 10 μ M was incubated in the presence of purified 1BKR-PRMT6-His 7.47 μ M, substrate 77.0 μ M and [^{3} H]-SAM 0.23 μ M respectively at 30°C. The methylated proteins were then bound to a phosphocellulose membrane, which was washed to remove any excess [^{3} H]-SAM, and the incorporation of the [^{3} H]-SAM quantified by scintillation counting.

The positive control in these experiments was the absence of inhibitors. The five compounds under evaluation showed weak or no inhibition activity towards 1BKR-PRMT6-His protein, while they appeared to inhibit CARM1 at a similar potency. Briefly, AzaSAM analogues **55** and **58**, which both contained 2-aminopyridine and either aminopropyl or amino acid side changes respectively showed very weak inhibition (approximately 40 % and 30 % inhibition at 100 μ M respectively). AzaSAM analogue **46** that targets the substrate-binding site with a terminal propylamine group was also a weak inhibitor when compared to activity against other enzymes. Similarly, either bisubstrate inhibitor **64** or lactam **39** showed very weak inhibition of the same protein (approximately 25 % and 20 % inhibition at 100 μ M, respectively (Figure 4-20).

These results are impressive when compared with the potent inhibition of CARM1 by the AzaSAM analogues **55**, **58**, **46**, and **64** reported earlier, which reflects a 90-fold selectivity of compounds for CARM1 versus PRMT6.

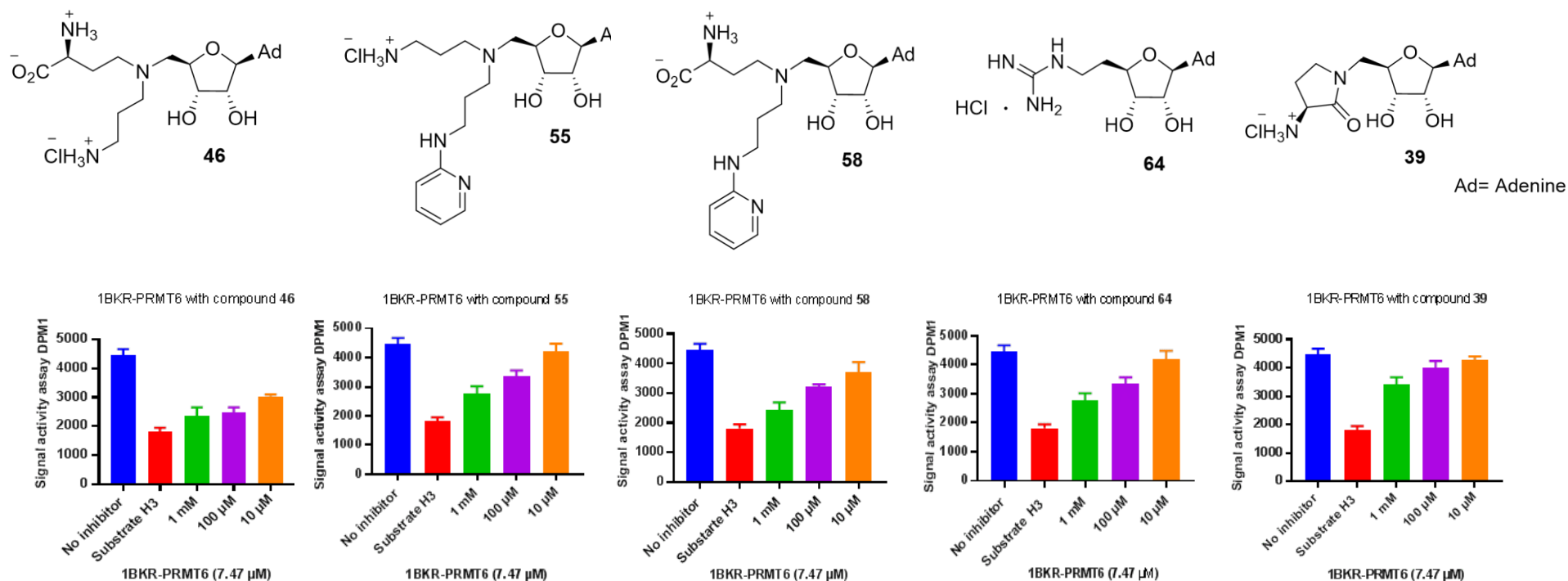


Figure 4-20: Preliminary activity test of inhibition against the catalytic activity of 1BKR-PRMT6-His.

Five graphs show three different concentrations of bisubstrate inhibitors 1 mM, 100 μ M and 10 μ M respectively, with 7.47 μ M of 1BKR-PRMT6-His and 77.0 μ M of the substrate. Positive control of 1BKR-PRMT6-His in the absence of bisubstrate inhibitors is highlighted in the blue and negative control of substrate without enzyme is highlighted in red. Error bars show the standard deviation of duplicate measurement at each inhibitor's concentration. DPM indicates the disintegrations per minute recorded by liquid scintillation counting and is proportional to histone methylation.

4.6. Crystallisation

After successfully producing and purifying 1BKR-PRMT6-His, attention turned to crystallisation of this construct. Solving the structure of PRMT6 linked with soluble tag bound to (SAH) as a model ligand has the potential to further our understanding of contact activation to a similar extent that solving the original PRMT6 structure did.

Attempts were made to co-crystallise 1BKR-PRMT6-His in the presence of SAH and the absence of SAH. Crystallisation trials were performed at various protein concentrations of 2.07, 3.00 and 3.20 mg/mL for 1BKR-PRMT6-His using JCSG+ and PACT screens. This concentration and screens were chosen based on the previous study when using the 1BKR soluble tag due to the presence of the tag increase the likelihood of forming crystals at these concentrations.^{94,103} This may suggest that 1BKR may be more likely to drive crystallisation when bound to a fusion partner. After three days, crystals were obtained in the presence of SAH with PACT screens in approximately 15 conditions (Table 4-2) and the morphology of obtained crystals are documented in Figure 4-21. Interestingly, no crystal growth was observed in all drops set up without SAH. This indicates that SAH, when bound to 1BKR-PRMT6, promoted protein stability and crystallisation.

Crystal Morphology	Crystallisation condition for PRMT6 with non-cleavable 1BKR soluble tag
Needles	PACT H1 0.2 M sodium fluoride 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT G1 0.2 M sodium fluoride 0.1 M Bis-Tris propane 7.5 20 % w/v PEG 3350 PACT H2 0.2 M sodium bromide 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT G2 0.2 M sodium bromide 0.1 M Bis-Tris propane 7.5 20 % w/v PEG 3350 PACT H4 0.2 M potassium thiocyanate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT G6 0.2 M sodium formate 0.1 M Bis-Tris propane 7.5 20 % w/v PEG 3350 PACT H6 0.2 M sodium formate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT H7 0.2 M sodium acetate trihydrate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT G8 0.2 M sodium sulfate 0.1 M Bis-Tris propane 7.5 20 % w/v PEG 3350 PACT H9 0.2 M potassium sodium tartrate tetrahydrate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT F9 0.2 M potassium sodium tartrate tetrahydrate 0.1 M Bis-Tris propane 6.5 20 % w/v PEG 3350 PACT F11 0.2 M sodium citrate tribasic dihydrate 0.1 M Bis-Tris propane 6.5 20 % w/v PEG 3350 PACT F12 0.2 M sodium malonate dibasic monohydrate 0.1 M Bis-Tris propane 6.5 20 % w/v PEG 3350 PACT E7 0.2 M sodium acetate trihydrate 20 % w/v PEG 3350 PACT E8 0.2 M sodium sulfate 20 % w/v PEG 3350 PACT F2 0.2 M sodium bromide 0.1 M Bis-Tris propane 6.5 20 % w/v PEG 3350

Table 4-2: conditions of 1BKR-PRMT6-His + SAH co-crystallisation.

Summary of co-crystallisation condition obtained for 1BKR-PRMT6-His + SAH using the composition of PACT premier (Molecular Dimensions) screen solutions.

Several crystals of 1BKR-PRMT6-His were obtained from a variety of conditions and were analysed using Diamond Light Source (DLS) beamline I24 (Figure 4-21). Datasets of crystals were collected, but the diffraction of 9-15 Å resolution was lower than appropriate to resolve the protein structure.

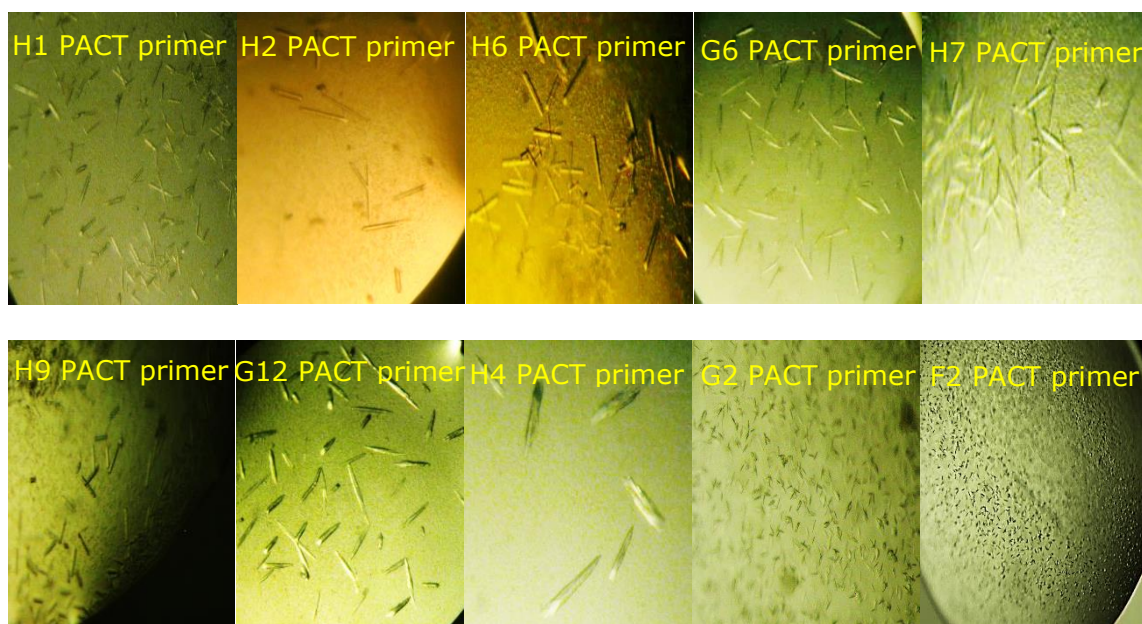


Figure 4-21: Images of crystals obtained from PACT primer (Molecular Dimensions) screen. Crystal obtained at 3 mg/mL in various crystallisation condition at a 1:3 ratio with SAH. These crystals were taken to Diamond Light Source for X-Ray diffraction experiments.

4.6.1. Data collection evolution of diffraction

Consequently, a further crystallisation step was attempted to optimise several conditions of PACT primer screen to yield further larger crystals by using a small 24 well screen. Crystals from that condition were collected and mounted for the X-ray home source. Unfortunately, similarly, poor resolutions were achieved for 1BKR-PRMT6-His crystals in complex with SAH.

Data was collected at a resolution of 9-15 Å, from more than three different translations on the crystal, but it was not possible to solve the structure of the protein complex. It was essential to confirm whether the crystals of 1BKR-PRMT6-His were of protein or salt. The results (Figure 4-22), indicate that the sample image of diffraction was protein based on the diffraction pattern.

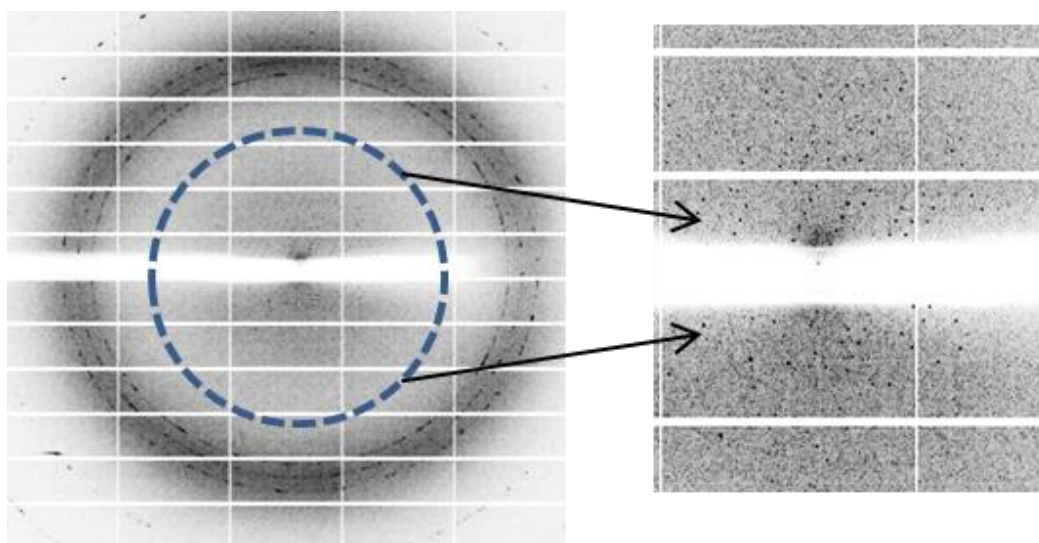


Figure 4-22: Diffraction pattern of crystal from H1 PACT Primer.

Diffraction image obtained from the crystal that grew in the crystallisation condition H1 PACT Primer, where the diffraction extended to a resolution of 9 Å.

A further important implication was to confirm whether the crystals were of 1BKR alone or 1BKR-PRMT6-His because it was possible that 1BKR or PRMT6 could be present as degradation products from the original purification during crystallisation (Figure 4-23). The crystals from condition H6 PACT screen were fished from crystallisation trials and solubilised in SDS-PAGE sample buffer. The solubilised crystals were loaded onto a gel along with control frozen samples of the proteins 1BKR-PRMT6-His and 1BKR. Analysis revealed that the band of crystals corresponded to the same molecular weight (56.0 kDa) as 1BKR-PRMT6-His. Samples of the frozen 1BKR-PRMT6-His and 1BKR were used as reference samples and confirmed that the crystals of protein had not degraded.

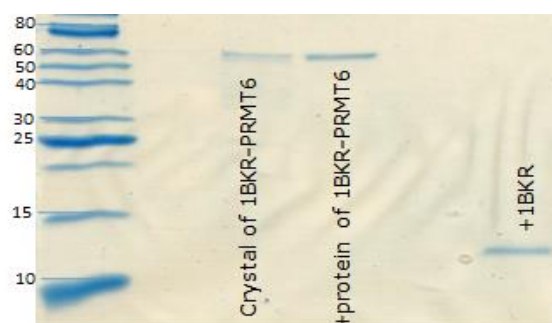


Figure 4-23: SDS-PAGE analysis to test the crystal of 1BKR-PRMT6.

SDS-PAGE analysis of 1BKR-PRMT6-His samples from crystallisation drops alongside control samples to assess the levels of degradation. Lane 1: protein ladder, lane 2: 1BKR-PRMT6-His sample as crystal from H6 PACT screen crystallisation drop 1, lane 3: 1BKR-PRMT6-His reference sample as a positive control, lane 4: 1BKR soluble tag reference sample as a positive control.

Data from all the crystals analysed by diffraction experiments are shown in Table 4-2. Four out of the eight crystals exposed to X-rays showed poor diffraction resolutions but confirmed the protein nature of these crystals. This is a common feature for many protein crystals and can be due to the high solvent content within the crystal and may also indicate a poor order of molecules.¹¹² There are a number of techniques that can be performed to improve diffraction resolution, including seeding, which involves using a crystal as a nucleation point for the growth of larger crystals¹¹³ and crystal dehydration, which is used to reduce the solvent volume and improve crystal packing.¹¹⁴ These techniques may be carried out in further future work.

Despite not solving the crystal structure of PRMT6 with the non-cleavable 1BKR soluble tag in complex with SAH ligand, the work documented in this chapter has narrowed down possible crystallisation conditions for future trials. Moreover, it may be worth using these conditions as templates for future optimisation experiments focused on improving the diffraction resolution.

4.7. Summary

Two His-tagged constructs of hPRMT6 were produced and purified in *E. coli*; however, both fusion proteins were present in an aggregated form in solution. Further attempts were focused on improving the solubility of the protein by using three solubility tags and as a consequence, produce higher titres of soluble protein. Initially, the MBP tag (41.0 kDa) was chosen⁹¹. However, the solubility of the protein was not improved when using this tag. Small tags linked to the N-terminus of the protein were also investigated, presence of 2GKG (14.0 kDa) (PDB code 2GKG) did not improve solubility. Finally, a 1BKR soluble tag (PDB code 1BKR) and use of Arctic Express cells improved protein solubility by co-expressing it with a chaperonin. Chaperonin could be successfully separated from protein and product monomer in solution; however, low yields were observed. Use of autoinduction media successfully improved the yield. Human 1BKR-PRMT6-His domain has been purified as a monomer and put into crystallisation trials. Although many crystals were obtained, unfortunately, all of them diffracted at a resolution too poor for structure resolution. Activity tests of the protein *in vitro* methylation assay showed activity as the monomer protein.

5. Chapter 5 Characterisation of human PRMT2

5.1. Full-length and catalytic domain constructs of human PRMT2

This chapter details the production and purification of human PRMT2 to facilitate an evaluation of its enzymatic activity. Previous studies have been reported that PRMT2 possesses weak methyltransferase activity when compared to PRMT1, CARM1, and PRMT6, respectively, and its optimal substrate has yet to be determined.^{35,7,32} Due to weak enzyme activity and the lack of structural data, thus far no PRMT2 inhibitor has been studied.³¹ The overall goal was to produce and then crystallise either a full-length PRMT2 protein or a protein containing only the catalytic core domain of PRMT2. These crystal structures would help to understand how this enzyme interacts with substrates and inhibitory compounds. It was essential to pinpoint the key regions and residues within the active site responsible for the binding of PRMT2. FL-PRMT2 has two major domains: a N-terminal region which possesses a SH3 domain, whereas the C-terminal domain represents the catalytic domain of PRMT2 (Figure 5-1).⁷

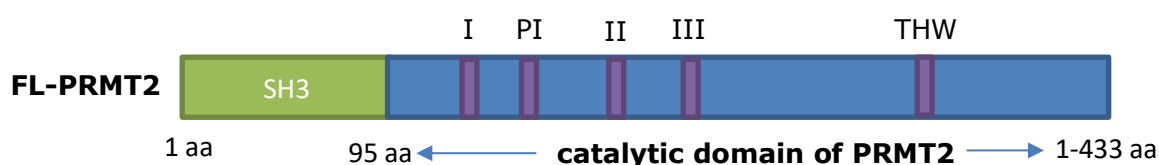


Figure 5-1: Overall domain architecture of full-length PRMT2.

The boundaries of each domain are indicated, SH3 domain is shown in olive green, and the catalytic core domain highlighted in blue, while the five motifs shown in purple.

There is no structure available for human PRMT2, but recently mouse (*Mus musculus*) and zebrafish (*D. rerio*) crystal structures of PRMT2 in complex with

reaction product, *S*-adenosylhomocysteine (SAH) have been produced.³¹ The catalytic core domain of the mouse PRMT2 domain structure (PDB code 5FUL)³¹ was solved to a resolution of 1.89 Å and the zebrafish PRMT2 domain (PDB code 5FUB)³¹ to 2.0 Å, respectively (Figure 5-2). The clustalw2 program was used to identify sequence similarities between zebrafish, murine and human PRMT2.³¹ It was found that human PRMT2 shared 80 % homology with mouse and 54 % with zebrafish PRMT2 and both mouse and zebrafish PRMT2 enzymes shared the same overall domain architecture.

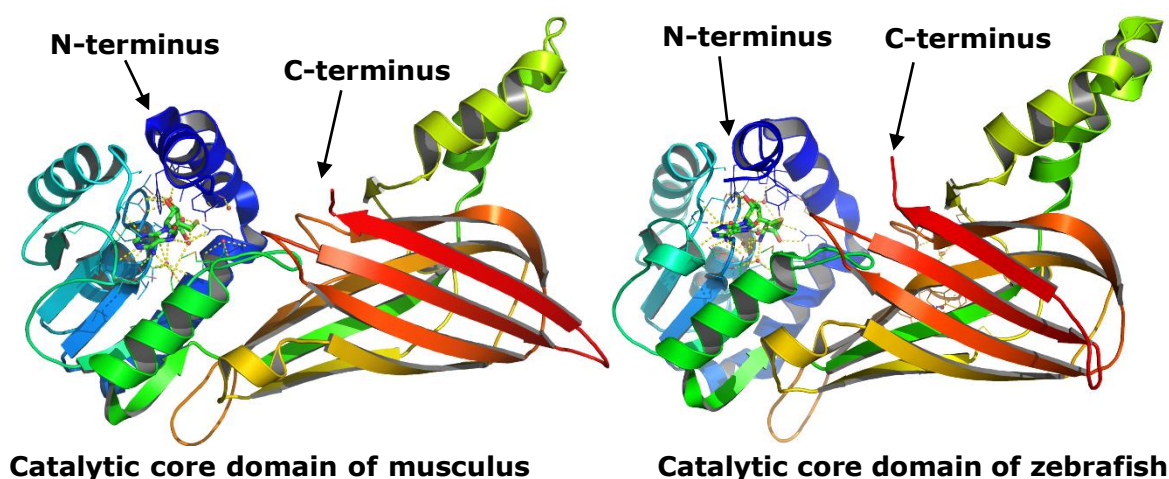


Figure 5-2: crystal structure of the mouse and zebrafish of PRMT2.

Overall domain architecture of either mouse (PDB code 5FUL)³¹ or zebrafish (PDB code 5FUB)³¹, which shows that these enzymes possessed a similar structure despite having different amino acid sequences.

Attention turned to aligning the homology model of human FL-PRMT2 structure with the crystal structure of mouse PRMT1 (PDB 1OR8)²⁴ to choose a construct suitable for producing a soluble protein. When the two structures were aligned, a loop was observed, which contained the majority of the first 95 residues that were unstructured (highlighted in blue, Figure 5-3). This loop connected to the catalytic core domain and could affect the solubility of the protein in solution. This

suggested that the first 95 residues could be removed, allowing the catalytic domain residues of PRMT2 to be produced as a soluble protein. Both full-length (FL) protein and residues 95-431 were chosen in term of producing a soluble protein.

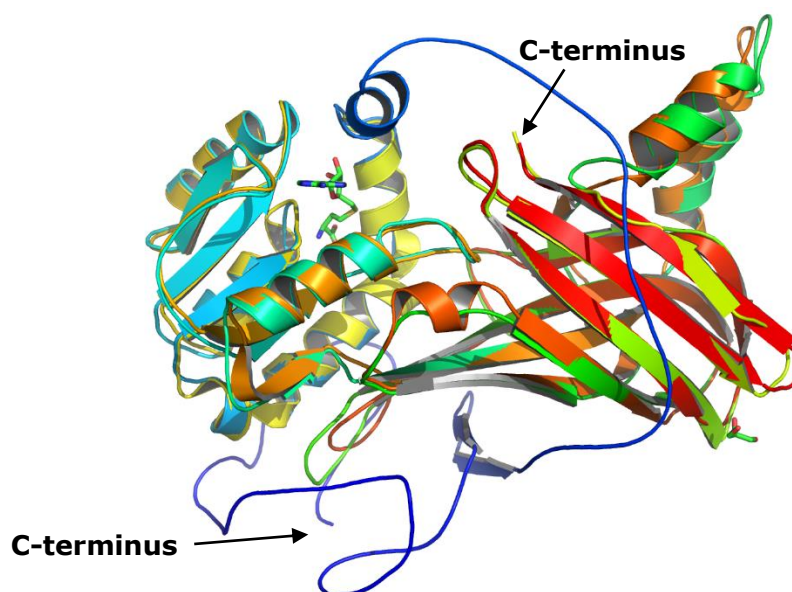


Figure 5-3: Alignment of human and mouse PRMT2 domains generated using PyMol.

Homology model of human of FL-PRMT2 (residues 1-431-6×His) was aligned with the catalytic domain crystal structure of mouse PRMT1 (PDB 1OR8).²⁴ Homology model structure of FL-PRMT2 showed that the first 95 residues are present in an unstructured region, which is highlighted in blue.

DNA corresponding to full-length PRMT2 (residues 1-431-6×His) and the shorter construct (residues 95-431-6×His) was amplified by PCR from the corresponding cDNA (Figure 5-4).

MATSGDCPRSESQGEEPAECSEAGLLQEGVQPEEFVAIADYAATDETQLSFL
 RGEKILILRQTTADWWGERAGCCGYIPANHVGHVDEYDPEDTWQDEEYF
 GSYGTLKLHLEMLA DQPRTTYKHSVILQNKESLTDKVILDVGC GTGIISLFCAH
 YARPRAVYAVEASEMAQHTGQLVLQNGFADIITVYQQKVEDVVLPEKVDVLVSE
 WMGTCLLFEFMIESILYARDAWLKEDGVIWPTMAALHLVPCSADKDYRSK VLF
 WDNAYEFNL SALKSLAVK EFFSKPKYNHILKPEDCLSEPCTILQLDMRTVQIS
 DLETLRGELRFDIRKAGTLHGFTAWFSVHFQSLQEGQPPQVLSTGPFHPTTH
 WKQTLFMMDDPVPVHTGDVVTG SVVLQRNPVWIRHMSVALSWAVTSRQDPTS
 QKVGEKVFP IWRHHHHHH

Figure 5-4: A sequence of FL-PRMT2 which is indicated in the red and blue.

Full-length PRMT2 sequences in which two domains are highlighted: SH3 domain omitted from the chosen sequence are in red, and the catalytic core domains are in blue.

5.1.1. Attempted production and purification of FL-PRMT2 construct

Full-length PRMT2 is challenging to purify using a bacterial expression system due to the significant potential for misfolding. Use of lower temperatures has been shown to slow translation and aid proper protein folding.¹¹⁵ The pCOLD expression vector was used for low-temperature protein production. This type of vector utilises a cold shock protein A (cspA) from *E. coli*, which can bind to RNA and act as a chaperone assisting with translation at low temperatures.¹¹⁵

Use of the pCOLD expression vector was explored in terms of improving FL-PRMT2 production. The FL-PRMT2 (residues 1-431-6×His) construct was successfully cloned into the pCOLD vector¹¹⁵ and transformed into LOBSTR BL21 (DE3)-RIL cells.⁸⁴ However, the yield of 6.0 µg/L obtained from the first stage of the nickel affinity chromatography was not enough to continue with gel filtration; the presence of impurities also prevented further assays using protein produced from this vector (Figure 5-5). Further optimisation was not carried out with this construct. Attention turned to improving the yield of FL-PRMT2 using trigger factor fusion as outlined below.

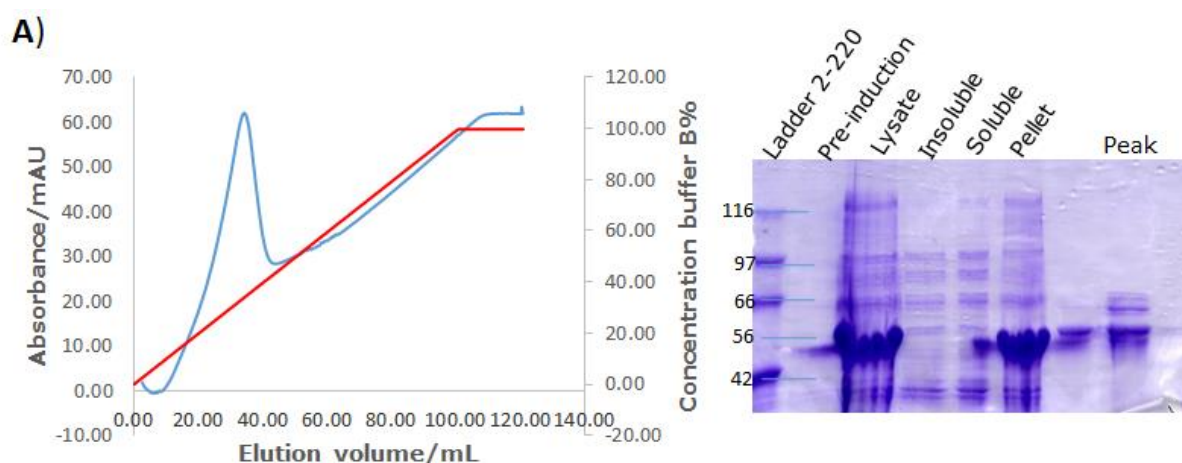


Figure 5-5: Optimisation of FL-PRMT2 using a pCOLD vector

A) UV trace of the nickel affinity chromatography of FL-PRMT2 (residues 1-431-6×His) produced in pCOLD vector. SDS-PAGE analysis of fractions: obtained from the nickel column; an insoluble sample from the cell pellet after centrifugation; the soluble fraction as loaded onto the nickel column; and the unbound sample obtained from the flow-through of the nickel column.

5.1.2. FL-PRMT2 production and purification in (TF) fusion protein

Zebrafish FL-PRMT2 (residues 2-408) has previously been crystallised.³¹ A plasmid encoding the full-length zebrafish sequence with an N-terminal GST tag in a pDEST20TM vector has been produced using the insect expression systems.³¹ Zebrafish of PRMT2 shares a 54 % homology with human PRMT2.

Production was attempted with full-length human PRMT2 (residues 1-431) using pCOLD encoding an additional trigger factor (49.0 kDa) for use as an N-terminus fusion tag.¹¹⁵ The trigger factor (TF) fusion tag is a 48.0 kDa prokaryotic chaperone protein,¹¹⁶ which promotes the correct folding of polypeptides as they are produced at the ribosome.¹¹⁶ It has been reported that trigger factor (TF) protein was produced well in *E. coli* and has been shown to increase the yield of the protein and improve the solubility of various target proteins when used as a fusion tag.¹¹⁶

Full-length PRMT2 (residues 1-431-6×His), including a 6×His tag at the C-terminus, was sub-cloned into the TF-pCOLD vector. Additionally, a PreScission protease cleavage site was added between the two proteins to allow separation of the fusion tag from the target protein. This construct was produced using protocols similar to those described earlier (Figure 5-5). The initial yield obtained from this purification was ~ 1.5 mg/L of bacterial culture, but most of the protein eluted in the void volume, indicating that the protein had aggregated. SDS-PAGE was carried out to determine the molecular weight of each protein band against a broad range protein marker (2-212 kDa) (Figure 5-6).

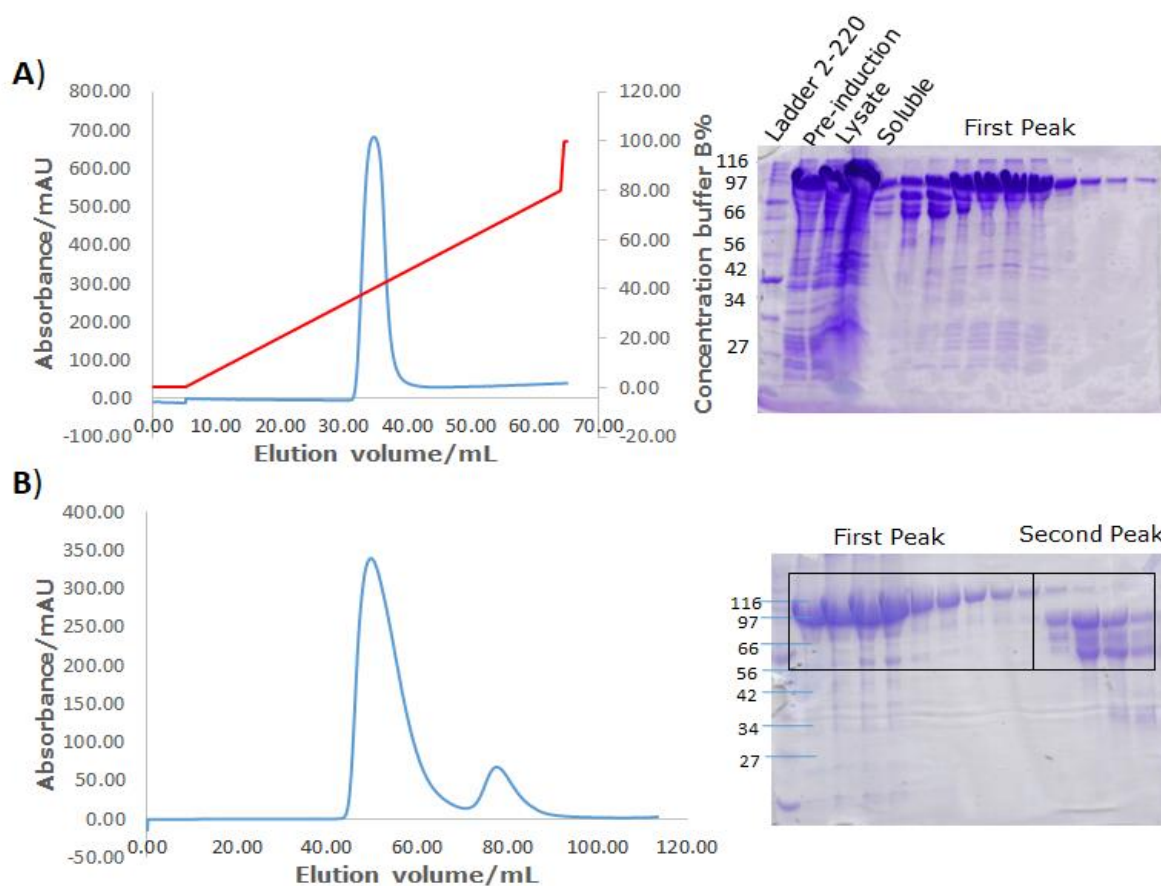


Figure 5-6: Optimisation of FL-PRMT2 with cleavable TF soluble tag.

A) UV trace of the nickel affinity chromatography of TF-PRMT2 (98.0 kDa) showing one peak which eluted at approximately 40 % of buffer B, with the corresponding SDS-PAGE gel. B) UV trace of gel filtration showed two main peaks eluted; the first corresponds to TF-PRMT2 with pure protein and the second to an impurity which eluted at approximately 80 mL, with the corresponding SDS-PAGE gel analysis.

5.1.2.1. Second purification of trigger factor-PRMT2 with cleavable linker

Attempts were made to remove the TF fusion tag using PreScission protease (HRV-C3 protease) from the full-length PRMT2. Once cleaved, as judged by SDS-PAGE (Figure 5-7B), a second gel filtration (Superdex™ 200 HiLoad 16/60) was used to separate cleaved FL-PRMT2 from the TF fusion tag. The result obtained from the gel filtration showed that FL-PRMT2 with the desired MW (48.0 kDa) had eluted into the void volume, which meant that the produced protein had become aggregated (Figure 5-7). By comparison, TF fusion tag was obtained as a stable monomer¹¹⁶ which eluted at volume 80 mL. Aggregation of FL-PRMT2 could have occurred within the cells while the protein was being produced (Figure 5-7). The resulting construct would most likely suffer solubility issues, meaning that full-length PRMT2 produced alone or with the soluble tag may not be the right approach.

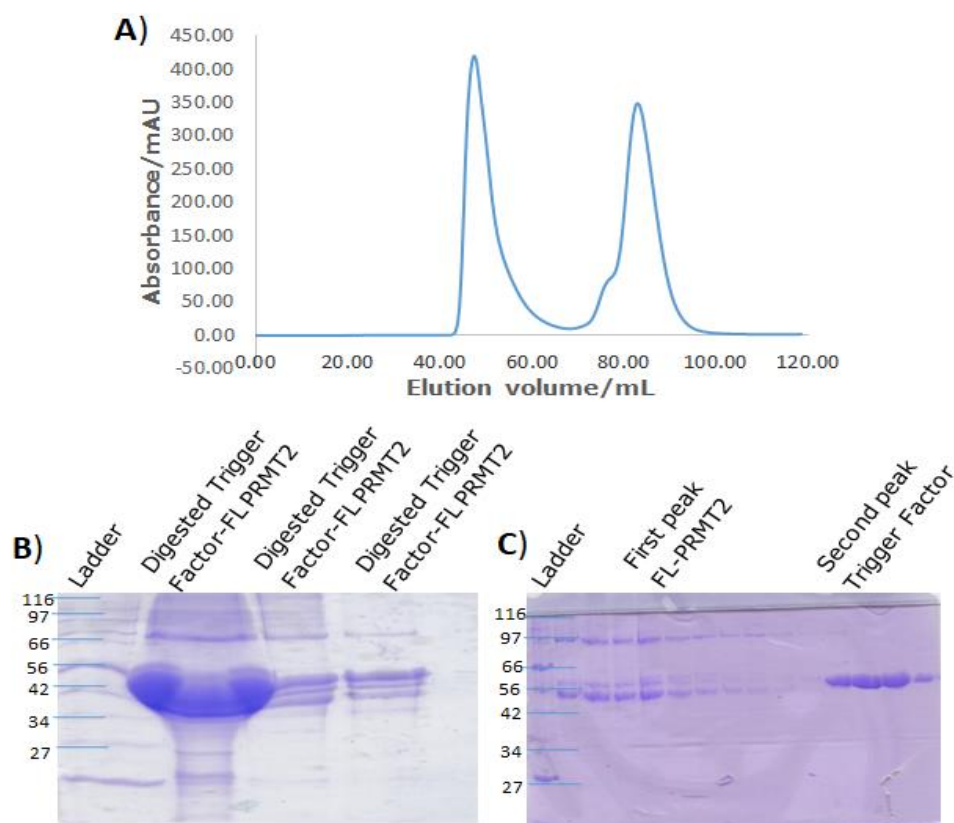


Figure 5-7: Second gel filtration used to separate cleaved FL-PRMT2 from TF tag.

A) UV trace of gel filtration using Superdex 200 HiLoad 16/60 for cleavable TF from PRMT2-His. B) SDS-PAGE showed the digestion of FL-PRMT2 and TF soluble tag using PreScission protease. Two bands were observed on the gel but were difficult to distinguish due to the similar molecular weights of FL-PRMT2 and TF soluble tag (49.0 and 48.0 kDa respectively). C) SDS-PAGE gel corresponding to the gel filtration column showed two main peaks; the first eluted was aggregation product at volume 45 mL, the second eluted monomeric soluble tag (TF) at a volume of approximately 80 mL.

5.2. Attempted production and purification of the catalytic core domain of PRMT2

5.2.1. PRMT2 construct design

As explained earlier, a mouse PRMT2 construct (mPRMT2 130-487) was successfully crystallised in complex with SAH but was determined to be inactive for methyltransferase activity *in vitro*.³¹ Therefore, the domain boundaries from D95 to W431 were chosen as the catalytic core domain for our PRMT2 construct. This construct was designed with the hope of producing a soluble protein that would allow us to assess the activity of the protein. This approach also allowed us to seek a crystal structure of PRMT2.

5.2.2 Purification of PRMT2 (residues 95-431-6×His) with soluble tags

In a further bid to obtain soluble protein, various constructs were generated with cleavable fusion tags; MBP-PRMT2, 2GKG-PRMT2 and 1BKR-PRMT2 with the aim of achieving the catalytic core domain (residues 95-431-6×His) as a soluble protein. Cleavable fusion tags were used to enhance the solubility or ease of purification of this construct. Initially, production and purification were attempted for both constructs with cleavable MBP and 2GKG tags; however, both showed a negative result indicating that the protein had aggregated in solution during the production of the protein (data not shown). This aggregation may be caused by the low T_m of the PRMT2 domains, where regions of the protein partially unfolded and interacted with other partially unfolded proteins. These regions of unfolding could lead to the conglomeration of many protein molecules as solid, fibrous precipitates. The data obtained from the limited proteolysis further supports the

theory that the N-terminus was partially unfolded. These observations also suggested that the poor solubility of PRMT2 during bacterial expression may be due to misfolding or aggregation.

5.2.3. PRMT2 in fusion with 1BKR soluble tag

Having failed to obtain a soluble protein with a MBP or 2GKG fusion protein, attention turned to producing and purifying PRMT2 as a 1BKR fusion protein. As was mentioned in the previous chapter, this tag had successfully produced soluble protein in the bacterial expression system (see Section 4.3.3. chapter 4.). Based on the prior work with PRMT6, it was predicted that if a 1BKR soluble tag was linked at the N-terminus of PRMT2 that could lead to an increase in expression levels and improve the stability of the protein. This construct (residues 95-431) was connected to the C-terminus of the 1BKR tag, with the *NotI* restriction enzyme used during cloning, use of this enzyme would produce a short linker between the two units, consisting of three alanine residues.

The catalytic domain of PRMT2 (residues 95-431) with the TEV cleavage site sequence was therefore sub-cloned into the 1BKR-pRSF13 vector. The expression plasmid was transformed into LOBSTR BL21 (DE3)-RIL cells for expression. Production and purification of this construct was performed in the same manner as 1BKR-PRMT6 (see Section 4.3.3.) using nickel affinity chromatography and gel filtration. After gel filtration on a HiLoad 16/16 column (Superdex™200 prep grad GE Healthcare) analysis revealed that two main peaks were observed, the first peak eluted at approximately 45 mL corresponding to the aggregated protein. A second more significant peak in terms of volume eluted at about 54 mL volume,

which corresponded to the soluble protein being present as an oligomer. The oligomer peak was broad with an estimated mass eight times greater than the predicted monomeric protein size; this data suggests that the protein is forming soluble micro-aggregates, or that the proteins are forming dynamic micelle-like structures with an equilibrium of eight units. Some PRMTs are known to form oligomers as a result of the association of dimers; these oligomers have been found in solution both *in vivo* and *in vitro*, respectively.¹⁰⁴ Previous research has indicated that the catalytic domain of mPRMT2 was present as a quaternary structure using small-angle X-ray scattering.³¹ Proteins from the two main peaks were examined on a 15 % (v/v) SDS-PAGE gel to confirm the presence of PRMT2 after purification (Figure 5-8). Highly pure protein was obtained, with a reduced yield of 0.01 mg/L.

A small amount of soluble protein obtained from this purification was measured for methyltransferase activity (Dr Emma Gunnell), but unfortunately, preliminary analysis showed that 1BKR-PRMT2 when present as an oligomer was not active (see Section 5.3.1.).

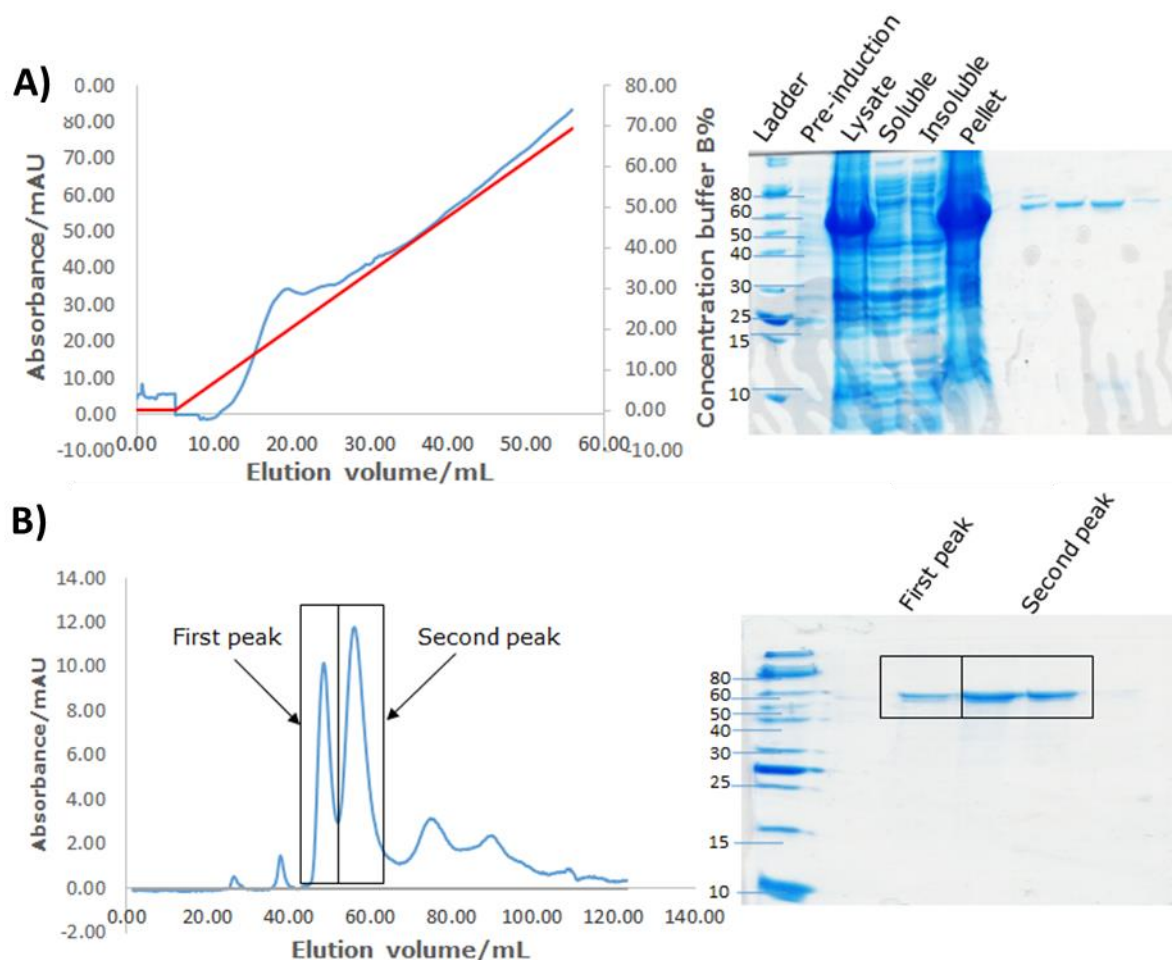


Figure 5-8: Purification of 1BKR-PRMT2-His (residues 95-431, MW=53.0 kDa).

A) Chromatogram of eluent from the nickel column showing a single peak containing 1BKR-PRMT2-His (residues 95-431) construct at 25 % of buffer B eluted with the corresponding SDS-PAGE gel. B) The gel filtration shows two main peaks with the corresponding SDS-PAGE gel; the first peak was eluted in the void peak and the second slightly larger peak trailing at 55.5 mL as oligomer protein (another two tiny peaks were identified as degradation protein).

Interestingly, this 1BKR soluble tag has been shown to enhance the correct folding of both PRMT2 and PRMT6 when used as a fusion tag. In future investigations, it might be possible to use this protein as a fusion tag to obtain proteins that otherwise tend to aggregate.

5.2.4. Analysis of 1BKR-PRMT2 by cleavage time-courses

To obtain further information about 1BKR-PRMT2 and its behaviour in solution, a cleavage assay was performed to ensure that the 1BKR tag was removed from PRMT2. To address this possibility, attempts were made at producing two constructs: one with the soluble protein corresponding to 1BKR-PRMT2 construct and one with the aggregated protein corresponding to 2GKG-PRMT2 construct for comparison using TEV protease.

These were performed with various concentrations of TEV protease to determine the suitable concentrations required for removal of the tag from the protein (Figure 5-9). Analysis of 1BKR-PRMT2 and 2GKG-PRMT2 cleavage time-course was attempted. Assays with 2GKG-PRMT2 construct showed all protease concentrations removed the 2GKG tag from the target protein, as determined by SDS-PAGE analysis (Figure 5-9B). Time-course analysis of 1BKR-PRMT2 cleavage revealed that only 10 % of the soluble tag had been removed from PRMT2 (Figure 5-9A). This finding was quite unexpected when compared to that for 2GKG-PRMT2. Interestingly, this result for 1BKR-PRMT2 is similar to that obtained for 1BKR-PRMT6 (see Section 4.3.3. chapter 4). It is not clear why the TEV protease failed to recognise the TEV cleavage site between 1BKR and PRMT2. It is also possible that some intermolecular interaction between 1BKR and the catalytic domain is facilitated due to the flexible nature of the linker, which may serve to protect the site from cleavage. The identity of these purified fractions was confirmed using western blot analysis.

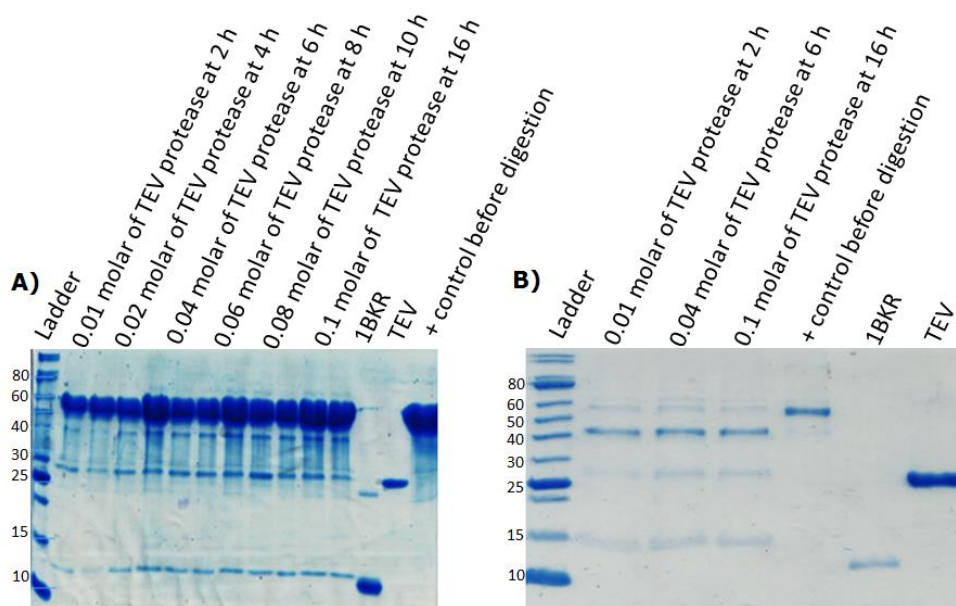


Figure 5-9: Time course cleavage of 1BKR-PRMT2 and 2GKG-PRMT2 (residues 95-431). A) SDS-PAGE gel showing a time course for a cleavage reaction at various concentration and time points using TEV protease to confirm cleavage of the 1BKR soluble tag from PRMT2. Three controls were used for comparison, which included TEV protease (27.0 kDa), 1BKR tag (13.0 kDa) and 1BKR-PRMT2 (53.0 kDa) before digestion. SDS-PAGE gel showed that 10 % of 1BKR had been cleaved from PRMT2. B) The SDS-PAGE gel corresponding to 2GKG-PRMT2 showed that most of the 2GKG tag had been removed from PRMT2.

5.2.5. Western blot

Further work was needed to determine whether peaks eluting during the gel filtration could be identified as either 1BKR-PRMT2 or 2GKG-PRMT2 or were another bacterial product. After production and purification of these two constructs, samples of the proteins were subject to SDS-PAGE which consisted of 1BKR-PRMT2-His (53.0 kDa) and 2GKG-PRMT2-His (54.0 kDa) in the presence of positive control PRMT1 (40.0 kDa) and a protein without His-tag (GST USHL1, 60.0 kDa) as a negative control. Western blot analysis was performed using the His-tag of 1BKR-PRMT2-His, 2GKG-PRMT2 and PRMT1 (Figure 5-10) and confirmed the presence of protein constructs carrying a His-tag. Nitrocellulose

stained with Ponceau S was run after ECL detection to compare the band of the western blot with protein marker (Figure 5-10).

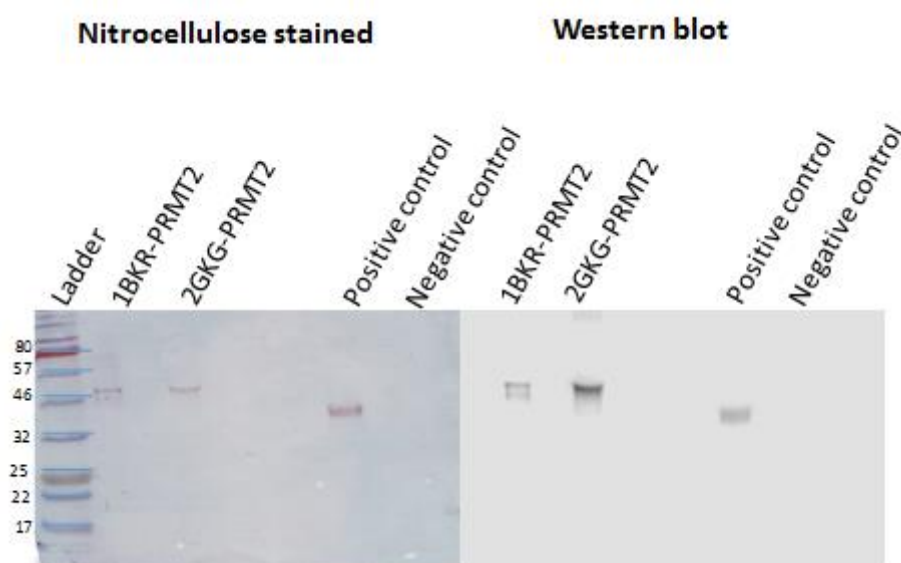


Figure 5-10: Western blot of both 1BKR-PRMT2 (53.0 kDa) and 2GKG-PRMT2 (53 kDa). A) ECL image detecting chemiluminescence shown on the right was used to identify the molecular weight of purified protein corresponding to 1BKR-PRMT2-His, 2GKG-PRMT2-His, respectively. Nitrocellulose stained with Ponceau stain after ECL detection consisting of 10 μ g His-tagged PRMT1 (MW = 40.0 kDa) was used as a positive control, and 8 μ g of protein GST USHL1 (MW = 60.0 kDa) was used as a negative control. B) 1BKR-PRMT2 and 2GKG-PRMT2 were identified by western blot using primary antibody (primary mouse anti-tetra His antibody) and secondary antibody (secondary anti-mouse horseradish peroxidase (HRP) conjugated antibody).

5.3. Purification of repressor splicing factor 1 (RSF1)

As discussed earlier, recombinant PRMT2 is described in various literature as having a weak activity for methylation of histones when compared to PRMT1, CARM1, and PRMT6 and that the optimal substrates for human PRMT2 are still unknown.³¹ In 2009, Lakowski, *et al.*, found the activity of PRMT2 to be 800-fold less active than PRMT1 for methylation of H4.¹⁰⁸ Moreover, evidence from *in vivo*

studies suggests that PRMT2 associates with β -catenin and has the ability to deposit an asymmetric dimethyl mark on histone 3 in arginine 8 (H3R8).¹¹⁹

Recently, repressor splicing factor 1 (RSF1) protein has been reported to co-purify with mouse mPRMT2, during expression in insect cells from *Spodoptera frugiperda* (Figure 5-11).³¹ This study shows that mPRMT2 can methylate arginine residues of RSF1, specifically monomethylation of arginine residues at position 24, 100, 111, and 139, and demethylation of residues at position 7 and 120 within a serine/arginine-rich (SR) region.³¹

Consequently, a synthetic peptide corresponding to RSF1 boundaries from N110 to R124 (highlighted in blue, Figure 5-11) was sought which could be used as a substrate with the aim of testing the activity of the purified 1BKR-PRMT2. This construct was chosen since two arginine residues (R111 and R 120) on this peptide having been shown to be methylated by a recombinant mPRMT2 with a GGGR motif.³¹ These residues, therefore, looked the most promising for the enzyme to methylate due to a large proportion of protein arginine methylation occurring within glycine/arginine-rich (GAR) motifs.²⁰

(A)

MSSGGT^{••}R₇VYVGGLVEGIKKEDLE[•]R₂₄EFDKYGKLNSVWV
ALNPPGFAFIEFENLQEAEDACSAMNGFEMLGATLKVEIS
RKRDGPRRGGGFRGGRGGNFRG[•]R₁₀₀GFGGPPGGGN[•]R₁₁₁S
FNSYGGGR^{••}R₁₂₀PYNRNNQRDEYGSGGGNFR[•]R₁₃₉SRSPLSRY

(B)

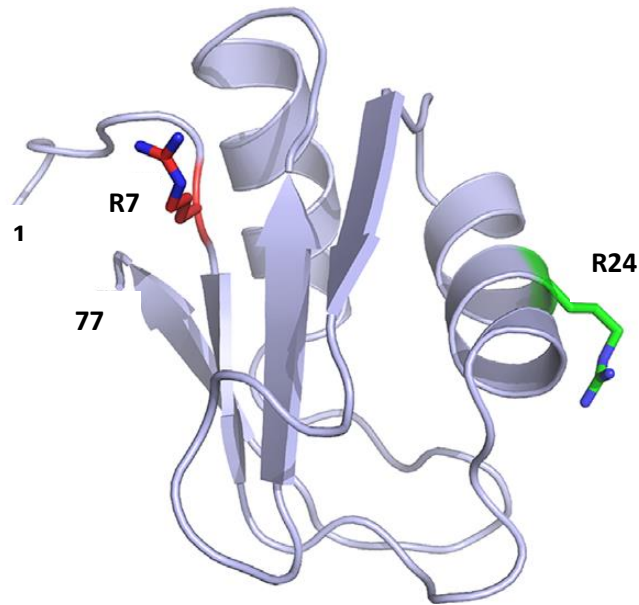


Figure 5-11: Model of the RRM domain indicating methylation sites on the RSF1 sequence.

A) A sequence representation of Repressor Splicing Factor 1 (RSF1) from *S. frugiperda* showing methylation sites. Monomethylation is highlighted in green and represented by one dot, whereas dimethylated arginines are shown in red with two dots. The sequence chosen to produce peptide (residues 110-124) for use as a substrate is highlighted in blue. (B) Homology 3D-Model of RSF1 RRM domain generated by Swiss-Model.¹²⁰ The arginine residues mono and demethylated are shown in the stick at the side chains.

As a complementary means for assessing the activity of 1BKR-PRMT2, several substrates peptides such as histone 3 and 4 (residues 1-22), and repressor splicing factor RSF1 (residues 110-124) were purified. These peptides and the splicing factor have been used for biochemical assays in several published studies,^{119,31} as described earlier.

With peptides of histones 3 and 4 in hand, we next sought to produce a peptide of RSF1 (residues 110-124) to use them as a substrate for measuring the activity of purified 1BKR-PRMT2. Brief optimisation of the radiometric (³H) activity assays were performed for 1BKR-PRMT2, as detailed in the following section.

A synthetic gene encoding RSF1 (residues 110-124) from *S. frugiperda* was successfully cloned into GB-pET21a and transformed into BL21-CodonPlus (DE3)-RIL competent cells (Agilent). As mentioned in the previous chapter 4, the GB tag was connected at the N-terminus of this peptide to aid the production of peptides and allow visualisation of the substrate using SDS-PAGE. Initial production of RSF1 was carried out in *E. coli* as a C-terminus His-tagged protein. Purification of the protein was performed in the same manner as previously, except the concentrated fraction from the nickel column was applied to HiLoad Superdex 75 prep grade (pg) column due to the small size of the non-tagged GB-RSF1 peptide (10.0 kDa). SDS-PAGE of corresponding fractions from the nickel affinity chromatography column and the gel filtration revealed that a peptide of the correct size was obtained. The final yield of soluble protein was 3.3 mg/L of bacterial culture (Figure 5-12).

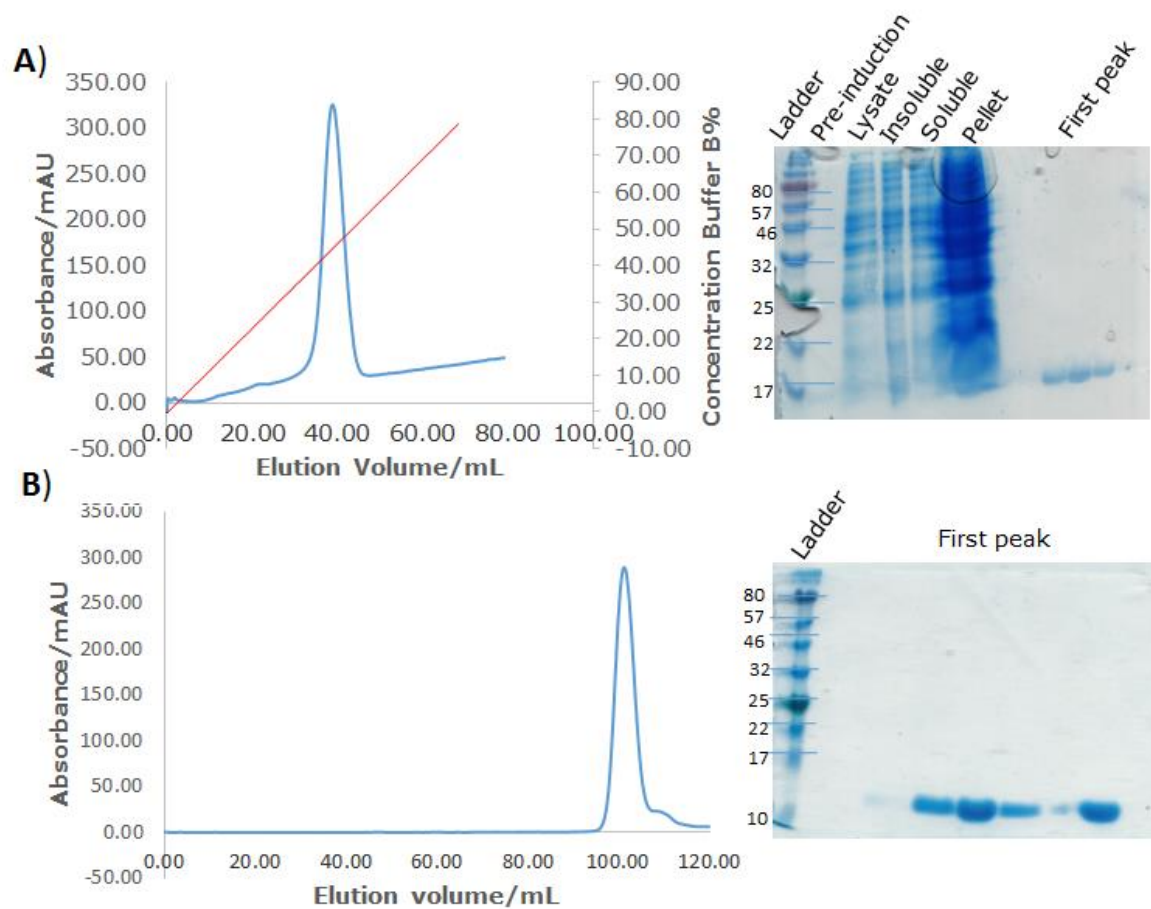


Figure 5-12: Purification of repressor splicing factor 1 (RSF1) (residues 110-124-6×His).

A) Nickel chromatogram shows a single peak containing the peptide of RSF1 (residues 110-124-6×His) with the corresponding SDS-PAGE gel. B) SDS-PAGE gel analysis of the eluent from the gel filtration, using Superdex S75 16/60, showed a single peak at 105 mL, suggesting that pure protein was obtained.

5.3.1. *In vitro* methylation assay

The methyltransferase activity of the soluble protein was evaluated by Dr Emma Gunnell. Peptides corresponding to histone H4 (residues 1–22), histone H3 (residues 1–22) and RSF1 (residues 110–124) were used as substrates to measure the activity of the 1BKR-PRMT2-His purified protein. The soluble proteins and substrates were prepared at the same concentration as used for PRMT1 assay, 27 μ M enzyme and 39 μ M substrate in the presence of 39 μ M SAM. The results and our initial findings suggested that poor incorporation of radioactivity was detected for 1BKR-PRMT2 with all peptides, H4 1–22, H3 1–22, or RSF1 110–124, in the presence of SAM (Figure 5-13). By comparison, the signal for methylation of H4 (residues 1– 22), by PRMT1 was much stronger (90000 DPM) (Figure 5-13).

As noted from several published papers,^{31,108} PRMT2 has been shown to have a weak activity towards peptides H4 (residues 1–22) and histone H3 (residues 1– 22). Results from our assays confirmed previous findings. A weak signal was also detected for 1BKR-PRMT2 with the peptide fragment from RSF1 (residues 110– 124). This was unexpected because mass spectrometry studies had shown that recombinant mPRMT2 was able to methylate arginine residues 111 (R111) on RSF1 when incubated with SAM.³¹ From these results, it could be concluded that 1BKR-PRMT2-His had weak methyltransferase activity when compared to that of PRMT1.

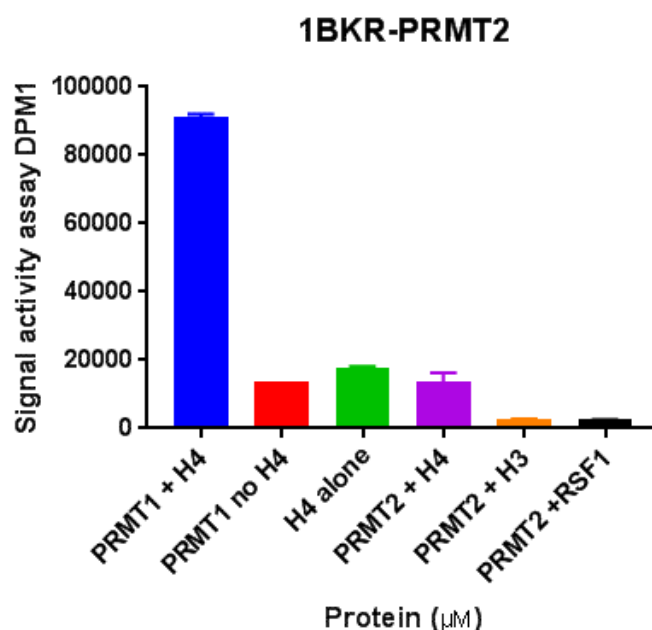


Figure 5-13: Initial activity test of 1BKR-PRMT2-His with three peptides

Three control reactions were conducted: PRMT1 with substrate H4 as a positive control, highlighted in blue; and negative controls represented by PRMT1 in the absence of histone peptide H4 indicated in red; and histone peptide H4 in the absence of PRMT1 shown in green. No significant increase in methyltransferase activity was detected for 1BKR-PRMT2-His in the presence of three peptides (H4, H3 and RSF1) when compared to the control sample of PRMT1 in the presence of substrate. Each line represents 1BKR-PRMT2-His in the presence substrate to confirm the activity of the enzyme. Error bars show the standard deviation from 2 repeats. DPM indicates the disintegrations per minute recorded by liquid scintillation counting and is proportional to histone and RSF1 methylation. Results plotted using GraphPad Prism.

5.3.2. Co-production of 1BKR-PRMT2-His with RSF1 or H3

There was no activity of 1BKR-PRMT2-His with peptides RSF1 (residues 110-124) and peptide histone tail H3 (residues 1-21) in the presence of SAM. Instead, co-purification of 1BKR-PRMT2-His with these peptides was performed with the hope that the resulting complex would be soluble. An increased chance of obtaining crystals of 1BKR-PRMT6-His may become possible if its complex with peptide is

stable. It was hypothesised that if the peptide was bound and spanned the active site of the protein, this domain could be involved in the recruitment of substrates to the catalytic domain. It was therefore of interest to explore this hypothesis because we became aware of a published study, which showed that mPRMT2 constructs methylated R111 of RSF1 peptide.³¹ Additionally, evidence from *in vivo* studies showed that PRMT2 associated with β -catenin to deposit an asymmetric dimethyl mark on histone 3 in arginine 8 (H3R8) of the target gene promoters.¹¹⁹

Initially, 1BKR-PRMT2-His (residues 95-371-6 $\times$ His) in a pRSF-13 vector and RSF1 peptide (residues 110-124) in a pET21a vector were co-produced in *E. coli* LOBSTR BL21 (DE3)-RIL cells with a view to detecting whether there is any interaction between the resulting proteins and peptides. Co-purification of the target protein with RSF1 peptide was performed in the same manner as previously, except the cells were lysed in low salt buffer comprising of 100 mM NaCl, 20 mM imidazole-HCl, 5 % (v/v) glycerol 50 mM Tris-HCl and pH 7.5 with the aim of enhancing the binding of the RSF1 peptide to 1BKR-PRMT2-His during purification. The two cell pellets were then mixed and co-purified together, initially by nickel affinity chromatography, and then by gel filtration.

As shown in the UV trace following gel filtration (Figure 5-14), protein eluted in four broad peaks at approximately 45, 55, 73 and 85 mL, respectively. The sample from each peak was loaded into the SDS-PAGE gel to determine the size and complex protein. SDS-PAGE analysis indicated that the second and third peaks eluting at 55 and 73 mL respectively contain both 1BKR-PRMT2-His (53.0 kDa) and peptide RSF1 (10.0 kDa), suggesting a complex of the two that persists during co-purification. A peak observed in the void volume at approximately 45 mL

corresponded to 1BKR-PRMT2-His solely, and the fourth peak at 85 mL contained contaminating proteins. These findings from the second and the third peak are in agreement with a report from Cura, *et al.*, (2016)³¹, which showed mouse PRMT2 co-produced with RSF1 in insect cells.³¹ The complex obtained was put into crystallisation trials at concentrations of approximately 3 mg/mL utilising PACT crystallisation screens, but unfortunately, no crystals of this complex were obtained.

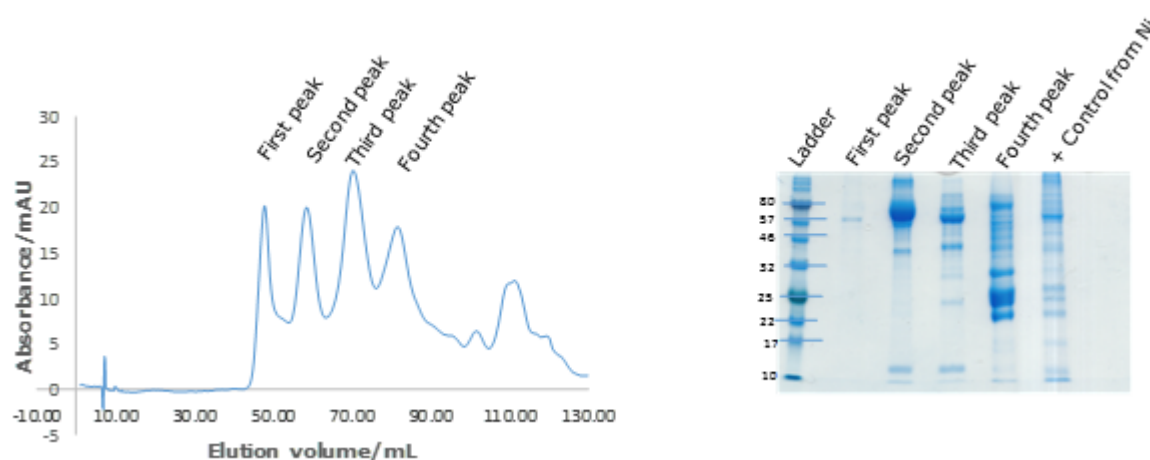


Figure 5-14: co-purification of 1BKR-PRMT2-His with peptide RSF1

UV trace of gel filtration chromatography of 1BKR-PRMT2-His co-purified with peptide RSF1 and subsequent SDS-PAGE analysis of the fractions. The fractions from each peak were pooled and concentrated separately and then loaded to the SDS-PAGE. 1BKR-PRMT2-His with RSF1 appeared to form a complex that persisted throughout purification to appear in fractions obtained at 65 mL (Peak 2) and 72 mL (peak 3), respectively. The dominant degradation species was obtained from fraction eluting at 84 mL (Peak 4), corresponding to a protein of ~ 25.0 kDa. Samples from the nickel column purification step were loaded onto SDS-PAGE for analysis. Nickel chromatograms and SDS-PAGE analysis of the nickel fractions are not shown.

Assays looking at co-purification of 1BKR-PRMT2-His constructs with peptide histone tail H3 (residues 1-21) were attempted. Low salt was used for this purification for the same reason as discussed for prior purifications.

Examination of the SDS-PAGE analysis of the fraction from the nickel column revealed bands corresponding to the molecular weight of 1BKR-PRMT2-His (53.0 kDa), the presence of high levels of protein contamination, as well as a band for the H3 peptide (10.0 kDa). All fractions containing either protein were combined and concentrated for gel filtration. As shown in the elution volume of gel filtration, four overlaps peaks were obtained (Figure 5-15). Each peak was examined by SDS-PAGE, and no complex peak was observed, indicating no detectable interaction between the two proteins, as both proteins eluted separately as an oligomer of 1BKR-PRMT2-His in the second peak and a definite shift in the elution volume of peptide H3 as a monomer.

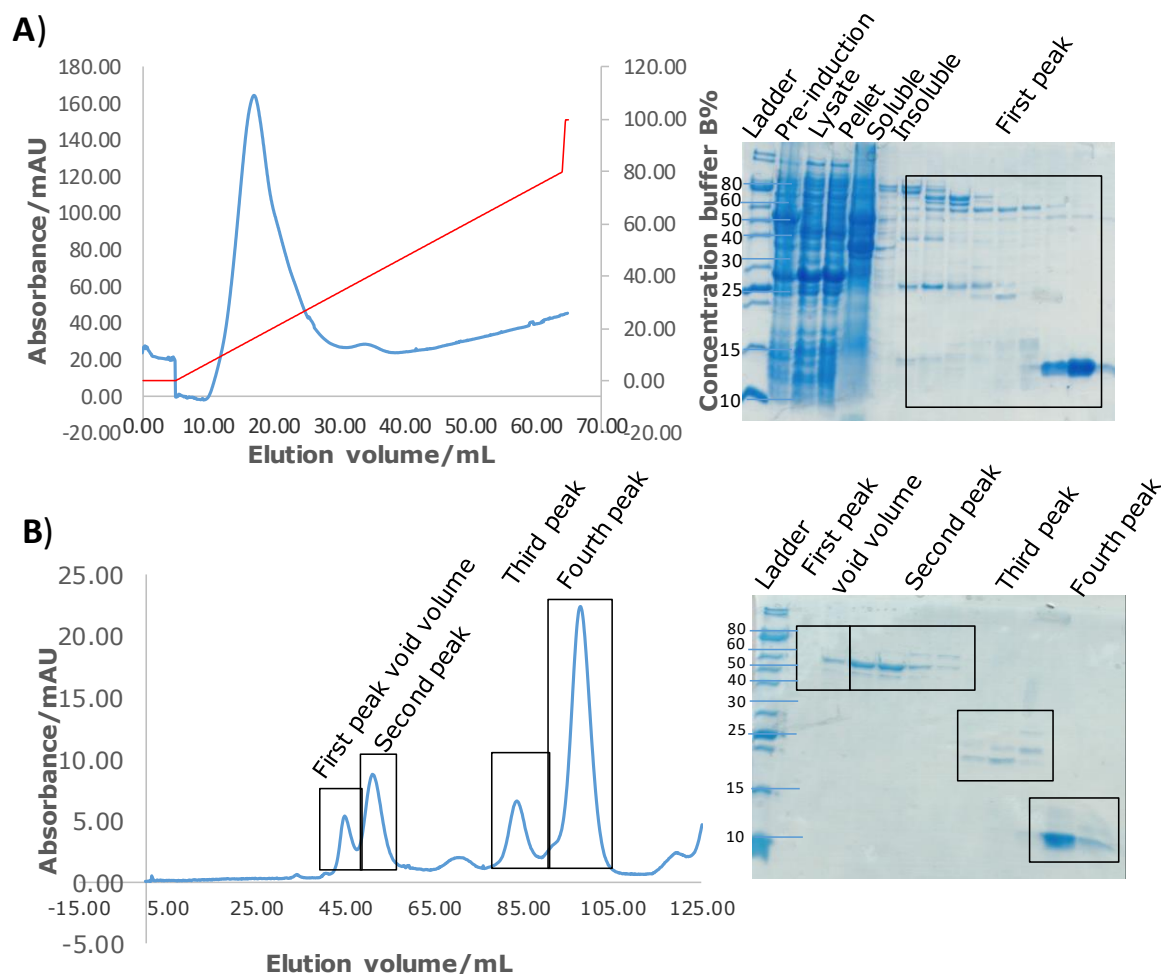


Figure 5-15: Co-purification of 1BKR-PRMT2-His with peptide H3 (residues 1-22).

A) UV trace of the nickel column chromatogram showed a co-production of 1BKR-PRMT2-His and peptide H3 eluting in one peak at approximately 40 % of buffer B. SDS-PAGE analysis showed two bands corresponding to 1BKR-PRMT2-His and peptide H3 as well as contamination protein. B) UV trace of eluent obtained from the gel filtration using Superdex 200 HiLoad 16/60. Aggregated and oligomeric 1BKR-PRMT2-His eluted at volume 45 and 55 mL respectively, peptide H3 eluted at 100 mL, and some contaminating proteins were observed at 80 mL, overloaded SDS-PAGE analysis of pooled fractions corresponded to 1BKR-PRMT2 and peptide H3 from the gel filtration.

5.4. Crystallisation of 1BKR-PRMT2-His

Crystallisation trials were carried out with the ambition of determining the structure of PRMT2 with 1BKR soluble tags. Attempts were made to co-crystallise the soluble protein of 1BKR-PRMT2-His in the presence and absence of the SAH ligand using the same conditions as used for CARM1 and PRMT6, respectively. Briefly, the protein was incubated in a 96-well plate at ambient temperature (19 °C) at various concentrations of 4.0, 3.0 and 2.5 mg/mL using JCSG+ and PACT screen conditions. The only precipitate was observed in wells containing only 1BKR-PRMT2 in the absence of SAH. However, crystals of different morphology (Figure 5-16) were formed after three days in both the PACT and JCSG+ screens in the presence of SAH under otherwise identical conditions (Table 5-1). Crystals were taken to Diamond light source (DLS) beamline I24 for data collection. The resolution obtained from data collection diffracted to 6-9 Å resolution, which was too poor to resolve the structure of the protein.

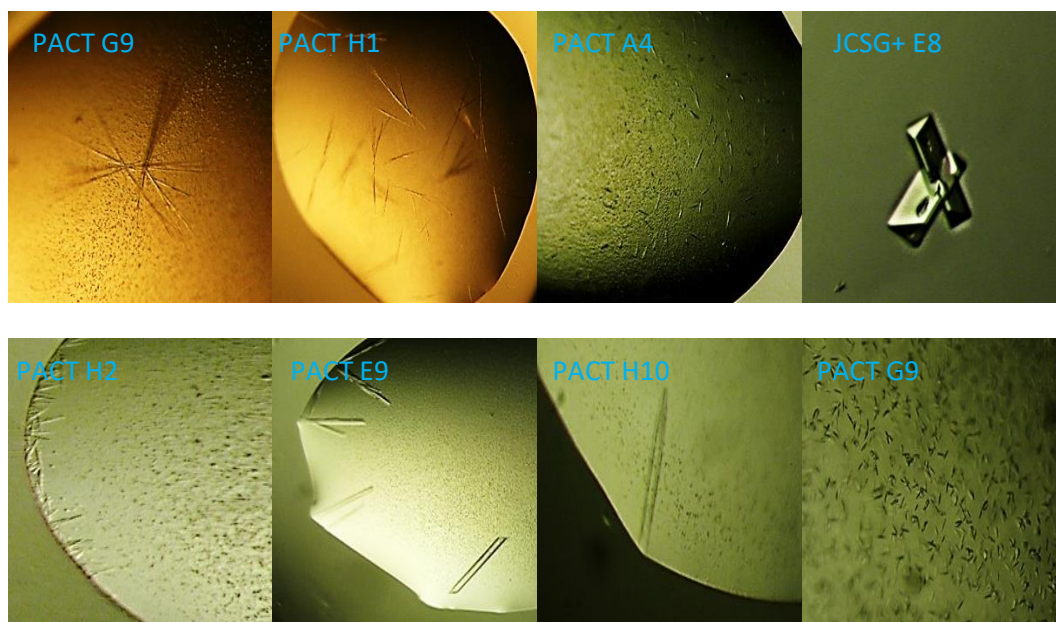


Figure 5-16: Crystals produced from the PACT screen for 1BKR-PRMT2-His with SAH. 1BKR-PRMT2-His crystals in complexes with SAH ligand were obtained from different conditions using a PACT screen.

Crystal Morphology	Crystallisation condition for 1BKR-PRMT2-His
Needles	PACT A4 0.1 M SPG 7.0 25 % w/v PEG 1500 PACT E9 0.2 M potassium sodium tartrate tetrahydrate 20 % w/v PEG 3350 PACT G9 0.2 M potassium sodium tartrate tetrahydrate 0.1 M Bis-Tris propane 7.5 20 % w/v PEG 3350 PACT H1 0.2 M sodium fluoride 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT H2 0.2 M sodium bromide 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT H3 0.2 M sodium iodide 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 PACT H10 0.02 M sodium/potassium phosphate 0.1 M Bis-Tris propane 8.5 20 % w/v PEG 3350 JCSG+ E8 1.0 M ammonium phosphate dibasic 0.1 M Sodium acetate 4.5 None

Table 5-1: Summary of crystallisation conditions where crystals were observed for 1BKR-PRMT2-His when in complex with SAH.

5.4.1. Data collection for PRMT2

The best resolution of 6-9 Å was not good enough, so the PACT screen was further optimised based on the conditions found in H1, H2, E9, and H9, respectively. This was performed in a 24 well crystallisation plate using various PEG 3350 concentrations (21-26 %) with a protein concentration of 3.0 and 3.30 mg/mL, and pH was varied from pH 7.0 to 9.0.

Large crystals were obtained after 3-6 days from various conditions. Several crystals from the H1 and H6 conditions were taken to Diamond Light Source (DLS) beamline I24 for data collection. Selection of these crystals were cryo-cooled using a selection of cryo-protectants (mother liquor with 25 % (v/v) glycerol either 25% (v/v) ethylene glycol). Again, similar resolutions between 6-15 Å were obtained from these crystals.

5.4.2. SDS-PAGE analysis western blot of a 1BKR-PRMT2-His crystal

It remained unclear as to whether the crystalline material at molecular weight 53.0 kDa was indeed 1BKR-PRMT2-His or some other protein artefact from the expression system. Western blotting was carried out, which confirmed the presence of the His-tag of protein attached to PRMT2 (Figure 5-17).

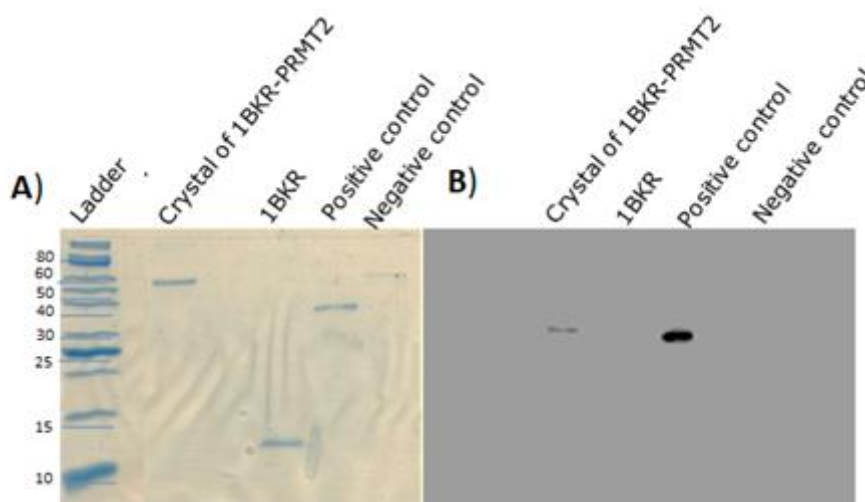


Figure 5-17: SDS PAGE analysis of 1BKR-PRMT2-His sample from crystallisation drop.

A) SDS PAGE gel shows. lane 1- 1BKR-PRMT2-His crystals combined from a selection of wells from H2 PACT screen (0.2 M sodium bromide, 0.1 M Bis-Tris propane, pH 8.5, 20 % w/v PEG 3350)1BKR-PRMT2-His crystals; lane 2- 1BKR soluble tag as a positive control; lane 3 and 4- sample with His tag (PRMT1 40.0 kDa) as a positive and sample GST USHL1 (MW = 60.0 kDa) without His tag as a negative control for the western blot respectively.

B) Western blot was carried out for 1BKR-PRMT2-His crystals with PRMT1 (MW = 40.0 kDa) containing a His-tag as a positive control and GST USHL1 protein (MW = 60.0 kDa) without a His-tag as a negative control, using primary antibody (primary mouse anti-tetra-His antibody) and secondary antibody (secondary anti-mouse horseradish peroxidase (HRP) conjugated antibody).

In this study, it has been shown that the presence of an IBKR soluble tag improved the expression and subsequent folding of recombinant proteins PRMT2 and PRMT6, respectively. This is, of course, a limited study; however, it shows the potential of using this tag for improving protein expression, crystallisation and enzymatic studies. The suitability of IBKR will need to be explored further in terms of affinity, size, cost and buffer conditions, all of which are important considerations when choosing a fusion partner for a protein. For example, production levels would need to be determined as a structural analysis requires higher protein concentrations that are typically associated with a more significant fusion tag; however, protein-protein interactions require lower production levels and therefore use of smaller tags.¹¹⁷ Whilst, IBKR worked well with PRMT2 and PRMT6, fusion partners do not always perform well with all target proteins, and the precise mechanisms for enhancing the solubility of a partner protein remains unclear. IBKR will need to be extensively investigated before its wide-spread acceptance as a soluble tag. Solubility tags are often used for affinity studies such as GST's or small ubiquitin containing tags, these tags allow for conventional purification, and it would be interesting to see if IBKR could be used in this manner. Additionally, some of these tags (GST, FLAG) are often used in tandem affinity purification,¹¹⁸ this approach is extremely useful in studying protein interactions.

5.5. Summary

At the first stage, we attempted to express and purify FL-PRMT2 alone or with a TF soluble tag using a pCOLD vector. Unfortunately, the majority of the protein eluted in the void volume, which meant that the protein had become aggregated. Attempts to produce and purify the catalytic domain of PRMT2 with cleavable MBP or 2GKG soluble tag failed to produce a soluble protein in solution. Also produced and purified was the catalytic domain of PRMT2 with the 1BKR soluble tag, which gave promising results and a soluble protein. Unfortunately, assays looking into the activity of the soluble protein in the presence of SAM with different substrates such as H3, H4 and RSF1 were negative as the protein proved to be inactive (Figure 5-13). Finally, crystallisation trials were carried out to attempt to crystallise soluble protein when complexed with the ligand SAH. Although crystals were successfully obtained, they all diffracted with a low resolution of approximately 6-15 Å, too poor to resolve the structure of the protein. Future experiments could be carried out using these conditions as templates for future optimisation experiments focused on improving the quality of the crystals. Such an effort would be justified since the results may provide an interesting insight into the structure of the active site of human PRMT2.

6. Chapter 6 Conclusion and future work

The initial objective of this work was to develop new bisubstrate inhibitors, which would be later evaluated as potential inhibitors of PRMTs. The first objective was the design and synthesis of AzaSAH analogue **46** connecting through a variable carbon chain, which was produced in good yields. The synthetic route involved Gabriel synthesis of protected 5'-amino-5'-deoxyadenosine **33**, successive reductive aminations with suitable aldehydes. Also attempted was the development of mimic compounds such as AzaSAH **55** and **58** analogues connecting to pyridine functionality. These analogues were not compatible with this route of reductive amination because the aldehydes could not be synthesised. Instead, they were successfully synthesised using an alkylation route. Temporary protection of the pyridine nitrogen by complexation with borane was essential in obtaining high overall yields. Deprotection of these compounds was successfully removing the Boc group by using neat TFA with water and purified using an Amberlite IRA 400 (Cl⁻) ion exchange resins to adjust the counter ion for biophysical characterisations. So, exploring the conditions for the synthesis of further substituted pyridine or guanidine is required in future to produce novel AzaSAM analogues, which may exhibit enhanced properties.

For the second objective, synthesis of SAH with butyl esters at the 2',3' hydroxyl **74** and **87** were successfully synthesised with the installation of the required thiol proceeding *via* a Mitsunobu transformation with thioacetic acid. These analogues were generated successfully using four steps, including a one-pot alkylation route.

An objective was to produce and purify hPRMT1, hCARM1 and then study the binding of the bisubstrate inhibitors with these enzymes using isothermal calorimetry. All the experiments were performed, and the results described in this thesis, provide the first insights into the overall structural arrangement of hPRMT1 and hCARM1. The crystallisation of hCARM1 in complex with bisubstrate inhibitors **55** and **58** was successful. The structure of CARM1 was solved by X-ray crystallography to 2.44 and 2.01 Å. This structure was the first to show that CARM1 when complexed with **55** or **58**, occupied both the arginine binding cavity and the surrounding pocket. The preliminary inhibition studies to measure selective inhibition between PRMT1 and CARM1 binding sites was carried out. The result indicated that *in vitro* compounds **55** and **58** had increased selectivity to CARM1 (IC₅₀ values of 0.18, 0.75 µM) over PRMT1 (IC₅₀ values of 52 µM and 9.7 µM), respectively.

To gain a greater understanding into the functionality of PRMT6, two constructs of hPRMT6 were chosen to be produced and purified in *E. coli*, both constructs expressed well, but unfortunately, purification was complicated by poor solubility in standard purification buffers. Optimisation of the protein by using various buffer conditions did not result in a soluble protein. It is expected that the protein may have been influenced by the formation of oligomers. It was attempted to improve the solubility of the protein by using three soluble tags with a long-term aim of obtaining soluble protein. Initially, a large MBP tag (41.0 kD) was chosen as a tag, as the presence of this tag has been shown to enhance protein solubilisation and protein crystallisation.⁹¹ Unfortunately the application of this tag failed to improved protein solubility. Expression and purification of PRMT6 was further optimised by

using a small linked 2GKG (14.0 kDa) tag at the N-terminus of the protein. However, the solubility of the protein was not improved when using this tag. Consequently, a pRSF13 vector which contained a 1BKR soluble tag at the N-terminus was chosen to improve the solubility of PRMT6 further. Using this optimisation, the soluble protein was obtained, which had been co-expressed with chaperonin; the importance of the chaperonin was investigated by mass spectrometry. As has been explained in chapter 4; chaperonin was successfully separated from the protein, and monomeric protein was produced but at low yield. This indicated 1BKR-PRMT6-His had been formed as a monomer in solution, and this was further confirmed by analytical *in vitro* methylation assay that showed the protein construct was catalytically active. Isothermal titration calorimetry was carried out to provide a more reliable means for investigating the binding of PRMT6 to SAH. However, as 1BK-PRMT6 domains appeared to be a stable product, these domains were set up in crystallisation trials. After three days, many crystals were formed using different buffer conditions, but unfortunately, low diffraction resolution were obtained.

Analysis of the enzymology of FL-PRMT2 (residues 1-431-6×His) and catalytic active domain of PRMT2 (residues 95-431-6×His) were performed. Initially, FL-PRMT2 (residues 1-431-6×His) construct was cloned, produced and purified using a pCOLD vector. Purification of this target protein was complicated by poor solubility when using standard purification buffers. To further investigate the solubility of FL-PRMT2 domains, attempts were made to produce and purify FL-PRMT2 with Trigger Factor soluble tag using the pCOLD vector. Unfortunately, the majority of the protein was eluted in the void volume, which meant that the protein had become aggregated.

The catalytic domain of PRMT2 (residues 95-341-6×His) was then produced as an MBP fusion protein with His tagging at the C-terminus. To assess the cleavage rate of the catalytic domain of PRMT2, one of the constructs possessed a TEV cleavage site between the soluble tag and a catalytic domain. Suitable conditions were determined for subsequent cleavage of the MBP tag and purification of PRMT2 (residues 95-431-6×His) from the resulting mixture. The result from both purifications was insoluble protein.

Further optimisation was attempted to produce and purify the protein with a 1BKR soluble tag. In this investigation, a soluble 1BKR-PRMT2-His protein was successfully produced. However, a 1BKR-PRMT2-His soluble protein was optimised by initially testing protein alongside various substrates such as H3, H4 and RSF1 concentrations in the presence of SAM, which provides the first step towards developing a rapid, accurate PRMT activity assay. The results of the correlational analysis of IBKR-PRMT2 showed that this enzyme was not active (Figure 5-13). Finally, crystallisation trials were carried out on soluble protein in complex with the ligand SAH. Crystallisation of this protein was successful when using different buffer conditions. Crystals were subject to X-rays at the synchrotron for data collection, and all the crystals were diffracted with the low resolution at approximately 6-15 Å, which was too poor to solve the structure. Future crystallisation attempts will focus on improving resolution and attempting to find the right substrate to test the activity of PRMT2.

Taken together, this data revealed essential insights into the structure of PRMT1 and CARM1 domains and highlighted the possible interaction of small molecule inhibitors with the catalytic activity of PRMT1 and CARM1. Overall, the structure

of human PRMT6 can be suggested where the bisubstrate inhibitors are likely to be involved in protein interactions, similarly to the homologues of PPMT1 and CARM1. Further investigation into improving the resolution of PRMT2 would be an exciting next step, and this would allow for a greater understanding of the enzymatic activity of PRMT2.

7. Chapter 7 Chemistry Experimental

Chemicals used for synthesis were purchased from Sigma-Aldrich and used without further purification unless otherwise stated. Acetone was supplied by Thermo Fisher Scientific. Dry THF was distilled from over sodium benzophenone, and dichloromethane was distilled from calcium hydride. Trifluoroacetic acid (TFA) was used as a neat liquid unless otherwise stated. Reactions were performed at ambient temperature (18-25 °C) under an argon atmosphere unless otherwise stated. Teflon-coated magnetic stirrer bars were used for stirring reactions. A Thermowatch-controlled DrySyn heating block was used to maintain reaction temperatures. Organic solutions were concentrated using a rotary evaporator with a water bath at 30 °C.

Thin-layer chromatography (TLC) was carried out using silica gel 60 F254 Merck pre-coated aluminium plates (0.2 mm thickness), with visualisation performed by UV light or exposure to potassium permanganate solution.

Flash chromatography was carried out on silica gel (particle size 35-70 µm, pore diameter 60 Å) with specified elution solvent system.

All ^1H NMR and ^{13}C NMR experiments were obtained using AV(III)500, DPX400, Bruker AV(III)400, DPX300 or JEOL EX270 NMR spectrometers at the field specified. Proton (δH) and carbon (δC) chemical shifts are quoted in part per million (ppm) and measured relative to a residual solvent such as the CDCl_3 7.27 ppm, D_2O 4.79 ppm and 2.54 ppm for $(\text{CD}_3)_2\text{SO}$, respectively. Assignments were made based on chemical shifts, coupling constants, COSY, DEPTQ, HSQC data and compared with spectra's of related compounds. The coupling constants (J) were

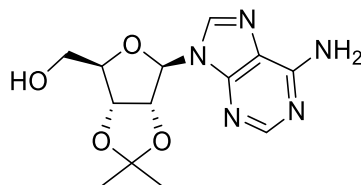
measured in Hertz and multiplicities quoted as singlet (s), doublet (d), double doublet (dd), triplet (t), quartet (q), quintet (quin), multiplet (m), broad (br) and apparent (app), or combination thereof.

High-resolution mass spectra were recorded using the open-access Mass Spectrometry Service at the University of Nottingham, School of Chemistry. Mass spectra were carried out using a MicroTOF mass spectrometer (electrospray ionisation ES, Bruker, US). Mass values (m/z) are reported in Daltons. High-resolution values were calculated to four decimal places from the molecular formula.¹⁰⁷

IR spectra were obtained for solid samples using AVATAR 320 FT-IR spectrometer (Thermo Scientific, Waltham, MA, USA) and for solution samples using Perkin Elmer FTIR 1600 (Thermo Scientific, Waltham, MA, USA). Optical rotations were obtained at ambient temperature using a Jasco DIP-digital polarimeter and ADP 440 polarimeter (A.KRÜSS, Germany). Uncorrected melting points were obtained using Stuart scientific SMP3 apparatus (SKS Science).

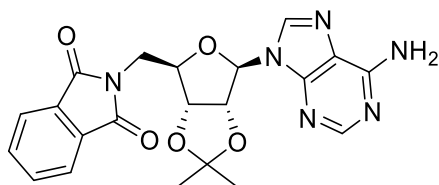
7.1. Experimental

Preparation of ((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methanol **31**⁵⁴



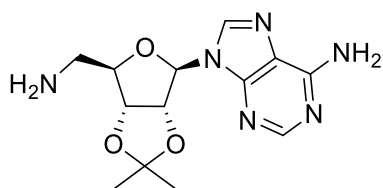
Adenosine (10.0 g, 37.4 mmol) and 2,2-dimethoxypropane (22.9 mL, 187.0 mmol) were dissolved in acetone (200 mL) at room temperature. *p*-Toluenesulfonic acid monohydrate (7.82 g, 41.2 mmol) was added slowly to the mixture. The mixture reaction was heated under reflux then immediately cooled to room temperature and left stirring for 16 hours. After that the mixture was quenched by slow addition of saturated aqueous sodium bicarbonate until pH rose to 6.0-7.0 and then extracted with chloroform (4 × 200 mL). The organic layers were combined and dried over magnesium sulfate, filtered, and concentrated under reduce pressure to afford the product as a white powder (9.88 g, 85 %): m.p. 216-218 °C [lit. 216 °C];⁵⁴ $[\alpha]^{24}_{\text{D}}$ -122.85 (c 1.0 in CHCl₃) [lit. $[\alpha]^{18}_{\text{D}}$ -65 in CHCl₃];⁵⁴ $\nu_{\text{max}}/\text{cm}^{-1}$ 3526 (O-H), 2995 (C-H), 1632 (C=N), 1113 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.33 (s, 1H, adenosine-*H*), 7.85 (s, 1H, adenosine-*H*), 5.95 (br. s, 2H, NH₂), 5.87 (d, *J*= 4.0, 1H, 1'-*H*), 5.23 (dd, *J*= 6.0, 2.0, 1H, 2'-*H*), 5.14 (dd, *J*= 6.0, 2.0, 1H, 3'-*H*), 4.56-4.53 (m, 1H, 4'-*H*), 4.01 (dd, *J*= 12.9, 1.7, 1H, 5'-CH_aH_b), 3.84 (dd, *J*= 12.9, 1.9, 1H, 5'-CH_aH_b), 1.66 (s, 3H, CH₃), 1.39 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 156.0 (C), 152.6 (CH), 148.4 (C), 140.3 (CH), 121.2 (C), 114.0 (C), 94.4 (CH), 86.1 (CH), 83.0 (CH), 81.7 (CH), 63.4 (CH₂), 27.7 (CH₃), 25.3 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd. C₁₃H₁₈N₅O₄⁺ 308.1353, found 308.1345.

Preparation of 2-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)isoindoline-1,3-dione **32⁵⁴**



A solution of diisopropylazodicarbonate (2.46 mL, 25.1 mmol) was added dropwise over 5 minutes to a mixture of 2',3'-*O,O*-isopropylideneadenosine **31** (3.56 g, 11.6 mmol), phthalimide (1.68 g, 12.2 mmol) and triphenylphosphine (3.21 g, 12.2 mmol) in anhydrous THF (100 mL). The mixture reaction was then left stirring for 16 hours at room temperature under argon. The resulting and precipitate was filtered and washed with diethyl ether to give the product as a white powder (4.50 g, 89 %): m.p. 151-156 °C [lit. 144-151 °C];⁵⁴ $[\alpha]^{24}_{\text{D}} -2.1$ (c 1.0 in CHCl₃) $[\alpha]^{18}_{\text{D}} -10.4$ in CHCl₃;² $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (N-H), 2984 (C-H, aromatic), 1717 and 1631 (C=O), 1079 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.08 (s, 1H, adenosine-*H*), 7.88 (s, 1H, adenosine-*H*), 7.82-7.78 (m, 2H, Ar-*H*), 7.73-7.69 (m, 2H, Ar-*H*), 6.06 (d, *J*= 1.8, 1'-*H*), 5.75 (b, s, 2H, NH₂), 5.56 (dd, *J*= 6.3, 1.8, 1H, 2'-*H*), 5.25 (dd, *J*= 6.3, 3.5, 1H, 3'-*H*), 4.57-4.54 (m, 1H, 4'-*H*), 4.08 (dd, *J*= 14.0, 6.0, 1H, 5'-*H*_a*H*_b), 4.00 (dd, *J*= 14.0, 6.0, 1H, 5'-*H*_a*H*_b), 1.61 (s, 3H, CH₃), 1.41 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 168.1 (C), 167.1 (C), 155.4 (C), 152.0 (CH), 150 (C), 136.4 (CH), 134.0 (CH), 134 (CH), 134.1 (C), 131.9 (C), 131.1 (C), 123.3 (CH), 122 (CH), 114.5 (C), 90.7 (CH), 85.1 (CH), 84.1 (CH), 82.6 (CH), 39.4 (CH₂), 27.1 (CH₃), 25.4 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₁H₂₁N₆O₅⁺ 437.1568, found 437.1551.

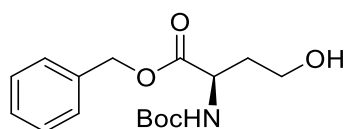
Preparation of 9-((3a*R*,4*R*,6*R*,6a*R*)-6-(aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine **33**⁵⁴



Hydrazine hydrate (4.15 mL, 132.3 mmol) was added dropwise over 5 minutes to a suspension of 5'-*N*-phthalimido-5'-deoxy-2'-3'-*O*-isopropylideneadenosine **32** (2.50 g, 5.73 mmol) in ethanol (60 mL). The mixture reaction was heated under reflux for 2 hours then warmed at room temperature and left stirred for 16 hours. The solution was then filtered and washed with ethanol. The filtrate was concentrated under reduced pressure and the residue dissolved in water (30 mL). The reaction was acidified using 1.0 M HCl until pH dropped to 1.0-3.0 and the resulting precipitate was filtered. Aqueous sodium hydroxide (1.0 M) was added slowly to the mixture filtrate until pH rose to 14.0, then the reaction mixture was extracted with chloroform (4 × 30 mL). The organic layers were combined and dried over magnesium sulfate, filtered, concentrated under reduce pressure and triturated with diethyl ether filtered to provide the product as white powder (1.46 g, 83 %): m.p. 204-205 °C [lit. 204-205 °C],⁵⁴ $[\alpha]^{24}_{\text{D}}$ -34.76 (c 1.0 in CHCl₃) [lit. $[\alpha]^{18}_{\text{D}}$ -32.5 (c 1.05 in CHCl₃)];⁵⁴ $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (N-H), 2992 (C-H), 1631 (C=N), 1078 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.34 (s, 1H, adenosine-*H*), 7.93 (s, 1H, adenosine-*H*), 6.10 (s, 2H, adenosine-NH₂), 6.05 (d, *J*= 3.1, 1H, 1'-*H*), 5.50 (dd, *J*= 6.5, 3.1, 1H, 2'-*H*), 5.04 (dd, *J*= 6.5, 3.4, 1H, 3'-*H*), 4.28-4.25 (m, 1H, 4'-*H*), 3.06 (dd, *J*= 13.3, 4.6, 1H, 5'-*H*_a*H*_b), 2.99 (dd, *J*= 13.3, 5.8, 1H, 5'-*H*_a*H*_b), 1.62 (s, 3H, CH₃), 1.39 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 155.7 (C), 153.1 (CH), 149.4 (C), 140.0 (CH), 120.4 (C), 114.6 (C), 90.7 (CH), 87.6 (CH), 83.6

(CH), 81.8 (CH), 43.9 (CH₂), 27.3 (CH₃), 25.4 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₁₃H₁₉N₆O₃⁺ 307.1513, found 307.1520.

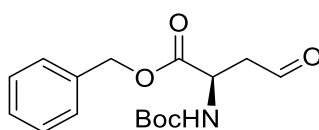
Preparation of benzyl(*tert*-butoxycarbonyl)-*D*- homoserinate **35**⁶⁶



To a stirred solution of 4-methyl morpholine (1.02 mL, 9.28 mmol) and Boc-Asp(OBzl)-OH (3.00 g, 9.28 mmol) was mixed in anhydrous THF (46 mL) at -15 °C. A solution of isobutyl chloroformate (1.20 mL, 9.28 mmol) was then added dropwise to the mixture reaction and left stirring for 10 min. After that, a solution of sodium borohydride (1.05 g, 27.8 mmol) was added in one portion, followed by dropwise addition (93 mL) of anhydrous methanol and the reaction left for another 10 min. A solution of 1.0 M HCl (18.6 mL) was added dropwise to the mixture reaction and left stirring for another one hour. The solvent was then concentrated under reduced pressure. The crude mixture was extracted with ethyl acetate (3 × 65 mL), and the combined organic layers were washed with 1.0 M HCl (1 × 37 mL), water (2 × 45 mL), 5 % NaHCO₃ (w/v) (1 × 46 mL), water (2 × 45 mL), then dried over Na₂SO₄. The organic layer was filtered, concentrated and purified by flash chromatography MeOH /CH₂Cl₂ (1/99: v/v) to give the compound was isolated as a clear, colourless viscous oil in a 90 % yield (2.70 g), [α]_D²¹ +13.43 (c 2.50 in CHCl₃); ν_{max} /cm⁻¹ 3626 (OH), 3432 (NH), 3008 (C-H), 1736 (C=O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 7.36 (m, 5H, *Ph*), 5.56 (s, 2H, CH₂Ph), 5.23 (s, 1H, OH), 5.20-5.16 (m, 1H, CHNH₂Boc), 4.64-4.60 (m, 2H, CH₂OH), 3.04 (dd, *J*= 17.0, 4.6, 1H, CH_aH_bCH), 2.86 (dd, *J*= 17.0, 4.8, 1H, CH_aH_bCH), 1.45 (s 9H, C(CH₃)₃); ¹³C

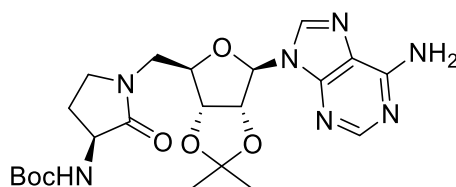
NMR (100MHz, CDCl₃) δ_c 172.2 (C), 156.3 (C), 135.2 (C), 128.7 (CH), 127.3 (CH), 126.9 (CH), 80.3 (C), 67.2 (CH₂), 59.2 (CH₂), 53.4 (CH), 35.7 (CH₂), 28.2 (CH₃); HRMS (ESI+) m/z : (M+Na⁺) calcd C₁₆H₂₃NO₅Na⁺ 332.1474, found 332.1467.

Preparation of benzyl (*R*)-2-((*tert*-butoxycarbonyl)amino)-4-oxobutanoate **36⁶⁷**



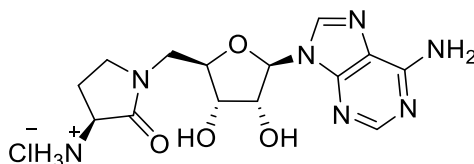
Dess-Martin periodinane (280 mg, 0.66 mmol) was added to a solution of alcohol **35** (70 mg, 0.22 mmol) in dry CH₂Cl₂ (2 mL) and left stirred for one hour. The mixture was quenched with saturated aqueous NaHCO₃ (10 mL) and the resulting mixture extracted with diethyl ether (3 × 10 mL). The organic layer was combined, dried over sodium sulfate, filtered, and the solvent was removed under reduced pressure to give colourless oil (60 mg, 86 %): $[\alpha]^{21}_D + 38.64$ (c 2.50 in CHCl₃); $\nu_{\max}/\text{cm}^{-1}$ 3437 (NH) 3009 (C-H), 1712 (C=O) 1161 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_H 9.73 (s, 1H, CHO) 7.36-7.32 (m, 5H, Ph), 5.19 (s, 2H, CH₂Ph), 4.66-4.63 (m, 1H, CHNHBoc), 3.07-3.03 (m, 2H, CH₂), 1.45 (s, 9H, C(CH₃)₃); ¹³C NMR (100MHz, CDCl₃) δ_c 205.1 (CH), 199.3 (C), 170.9 (C), 135.1 (C), 128.9 (CH), 128.5 (CH), 128.5 (CH), 80.31 (C), 67.5 (CH₂), 48.78 (CH), 45.9 (CH₂), 28.2 (CH₃); HRMS (ESI+) m/z : (M+Na⁺) calcd C₁₆H₂₁NO₅Na⁺ 330.1317, found 330.1307.

Preparation of *tert*-butyl((*S*)-1-(((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)-2-oxopyrrolidin-3-yl)carbamate **38**



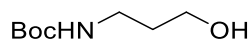
NaBH(OAc)₃ (47 mg, 0.22 mmol) was added to a suspension of aldehyde **36** (55 mg, 0.17 mmol) and a solution of secondary amine **33** (50 mg, 0.016 mmol) in 1, 2- dichloroethane (5 mL) at room temperature under argon. The reaction mixture was left stirring for four hours and then quenched with saturated aqueous Na₂CO₃ (5 mL). The mixture was then extracted with CH₂Cl₂ (2 × 10 mL) and dried over sodium sulfate, filtered, concentrated *in vacuo*. The residue was purified by silica gel chromatography using MeOH/CH₂Cl₂ (2/100, v/v) as eluent to provide the product as a colourless oil (0.04 g, 57 %): [α]_D²³ -23.68 (c 2.50 in CHCl₃); ν_{max} /cm⁻¹ 3413 (NH) 2986 (C-H), 1697 and 1631 (C=O) 1160 (C-O); ¹H NMR (400 MHz, CD₃OD) δ_{H} 8.34 (s, 1H, adenosine-*H*), 8.26 (s, 1H, adenosine-*H*), 6.19 (d, *J*= 2.8, 1H, 1'-*H*), 5.52-5.48 (m, 2H, CHNH_{Boc}), 5.50 (dd, *J*= 6.2, 3.0, 1H, 2'-*H*), 5.13 (dd, *J*= 6.2, 3.0, 1H, 3'- *H*), 4.38-4.34 (m, 1H, 4'-*H*), 4.13 (app. t, *J*= 6.0, 1H, CHNH), 3.65 (dd, *J*= 14.0, 5.5, 1H, 5'-*H*_a*H*_b), 3.61 (dd, *J*= 14.0, 5.5, 1H, 5'-*H*_a*H*_b), 3.33.3.30 (m, 2H, CH₂N), 2.39 (m, 1H, CH_a*H*_bCH) 1.99 (m, 1H, CH_a*H*_bCH) 1.60 (s, 3H, CH₃), 1.46 (s, 9H, C(CH₃)₃) 1.38 (s, 3H, CH₃); ¹³C NMR (100 MHz, CD₃OD) δ_{C} 172.9 (C), 171.1 (C), 155.8 (C), 153.2 (CH), 149.3 (C), 139.9 (CH), 120.1 (C), 114.8 (C), 90.0 (CH), 85.2 (CH), 83.5 (CH), 82.0 (CH), 79.9 (C), 52.3 (CH), 45.8 (CH₂), 45.0 (CH₂), 28.33 (CH₃), 27.7 (CH₂), 27.14 (CH₃), 25.3 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₂H₃₁N₇O₆⁺ 490.2409, found 490.2407.

Preparation of (S)-1-(((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-2-oxopyrrolidin-3-aminium chloride **39**



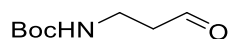
For deprotection of the compound **39**, protected tertiary amine **38** (40 mg, 0.81 mmol) was dissolved in neat TFA (5 mL) and left stirred for 1 hour at room temperature. Water (0.5 mL) was then added to the mixture and continued stirring for another 4 hours. The reaction mixture was concentrated and dissolved in H₂O (5 mL), and then washed with EtOAc (2 × 5 mL). The resulting of the aqueous layer was concentrated and subjected to Amberlite IRA 400 (Cl⁻) ion exchange resin, eluting with H₂O to give the product as pale yellow solid (0.02 g, 71 %): $[\alpha]^{23}_{\text{D}} +11.43$ (c 1.0 in H₂O) $\nu_{\text{max}}/\text{cm}^{-1}$ 3606 (O-H), 3413 (N-H), 3007 (C-H), 1631 (C=O) 1156 (C-O); ¹H NMR (400 MHz, D₂O) δ_{H} 8.40 (s, 1H, adenosine-*H*), 8.31 (s, 1H, adenosine-*H*), 6.03 (d, *J*= 4.0, 1H, 1'-*H*), 4.77 (dd, *J*= 6.5, 1.3, 1H, 2'-*H*), 4.38-4.34 (m, 2H, 3'-*H* and 4'-*H*), 4.13 (app. t, *J*= 5.1, 1H, CHNH₂), 3.65-3.63 (m, 2H, CH₂N), 3.49 (dd, *J*= 14.3, 4.8, 1H, 5'-*H_aH_b*), 3.45 (dd, *J*= 14.3, 5.9, 1H, 5'-*H_aH_b*), 2.48 (m, 1H, CH_{*a*}*H_b*CH) 1.99 (m, 1H, CH_{*a*}*H_b*CH); ¹³C NMR (100 MHz, D₂O) δ_{C} 172.9 (C), 171.1 (C), 155.8 (C), 153.2 (CH), 149.3 (C), 139.9 (CH), 90.0 (CH), 85.2 (CH), 83.5 (CH), 82.0 (CH), 52.3 (CH), 45.8 (CH₂), 29.7 (CH₂), 27.7 (CH₂); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₁₄H₂₀N₇O₄⁺ require 350.1571, found 350.1585.

Preparation of *tert*-butyl(3-hydroxypropyl)carbamate **41**⁶⁸



A solution of di-*tert*-butyl dicarbonate (9.60 g, 44.0 mmol) in dry CH₂Cl₂ (50 mL) was added slowly to the mixture solution of Et₃N (5.60 mL, 40.0 mmol) and 3-aminopropanol (3.00 g, 40.0 mmol) in dry CH₂Cl₂ (50 mL) at room temperature under argon. The suspension was heated under reflux for 16 hours and then cooled to room temperature. The mixture of this reaction was quenched with saturated aqueous NH₄Cl (30 mL) and then extracted with CH₂Cl₂ (3 × 50 mL). The organic layers were combined and washed with brine (1 × 20 mL). The organic solution was continued to dry over MgSO₄, filtered, and concentrated under reduced pressure. Column chromatography was used to purify the residue by using MeOH/CH₂Cl₂ (2/9, v/v) as eluent to give the product as a pale yellow oil (5.50 g, 78 %): $\nu_{\max}$ /cm⁻¹ 3600 (O-H), 3458 (N-H), 1695 (C=O), 1165 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 4.93 (s, 1H, NH), 3.68 (app. q, *J*= 5.7, 2H, CH₂OH), 3.28 (app. q, *J*= 6.2, 2H, CH₂NHBoc), 2.19 (t, *J*= 5.3, 1H, OH), 1.67-1.63 (m, 2H, CH₂CH₂CH₂), 1.45 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 157.1 (C), 79.5 (C), 59.2 (CH₂), 36.9 (CH₂), 32.8 (CH₂), 28.3 (CH₃); HRMS (ESI+) *m/z*: (M+Na⁺) calcd C₈H₁₇NNaO₃⁺ 198.1106, found 198.1105.

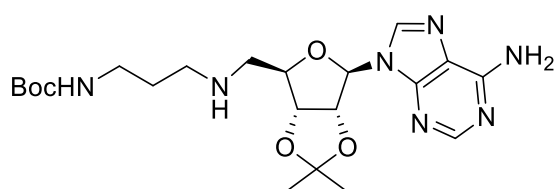
Preparation of *tert*-butyl(3-oxopropyl)carbamate **42**⁶⁹



The method according to Swern *et al.*³⁵ A solution of oxalyl chloride (3.92 mL, 46.4 mmol) was added to a solution of dimethyl sulfoxide (6.50 mL, 91.6 mmol) in dry dichloromethane (130 mL) at -78 °C and stirred for 30 min. Then, a

suspension of alcohol **41** (5.42 g, 30.95 mmol) in anhydrous CH₂Cl₂ (100 mL) was added dropwise over 5 minutes to the mixture at -78 °C and left stirring for 1 hour. Et₃N (20 mL, 143.6 mmol) was added dropwise to the mixture at -78 °C and stirred for another 30 minutes. The mixture was then warmed to 37 °C and quenched with 1.0 M HCl (100 mL). The combination was extracted with ethyl acetate (4 × 40 mL), and the organic layer was continually washed with saturated aqueous NaHCO₃ (75 mL). The organic layer was collected and dried over MgSO₄ filtered, then concentrated under reduce pressure to afford the product as pale yellow oil (4.56 g, 85 %): $\nu_{\text{max}}/\text{cm}^{-1}$ 3455 (N-H), 1699 (C=O), 1240 (C-N), 1159 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 9.70 (app. t, *J*= 1.6, 1H, CHO), 4.91 (s, 1H, NH), 3.36 (app. q, *J*= 6.4, 2H, CH₂NH), 2.60-2.56 (m, 2H, CH₂CH₂CH), 1.45 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 201.9 (CH), 155.9 (C), 78.3 (C), 43.8 (CH₂), 34.0 (CH₂), 27.6 (CH₃); HRMS (ESI+) *m/z*: (M+Na⁺) calcd C₈H₁₅NNaO₃⁺ 196.0950, found 196.0954.

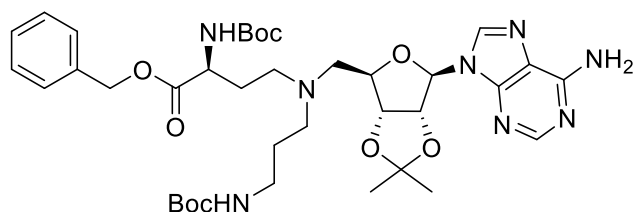
Preparation of *tert*-butyl-(3-((((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)amino)propyl)carbamate **43**⁵⁹



The suspension of aldehyde **42** (600 mg, 3.40 mmol) in 1, 2-dichloroethane (40 mL) was added to a suspension of amine **33** (1.16 g, 3.80 mmol) in 1,2-dichloroethane (40 mL) at 20 °C under argon. The mixture reaction was warmed

to 60 °C and then cooled to room temperature. NaBH(OAc)₃ (1.00 g, 4.80 mmol) was added slowly in several portions to the mixture and the reaction then left stirring for 16 hours at rt. The solution was quenched with saturated aqueous Na₂CO₃ (40 mL), and then extracted with CH₂Cl₂ (2 × 80 mL). The organic layers were combined and dried over Na₂SO₄, filtered, concentrated under reduce pressure then the residue was purified by silica gel column chromatography using MeOH/EtOAc (4/100, v/v) to give the product as a white foam (1.53 g, 66 %): m.p 55-57 °C, mp [lit. 55 °C];⁵⁹ [α]²⁴_D -13.1 (c 1.0 in CHCl₃) [lit. [α]²⁴_D -10.5 in CHCl₃];⁵⁹ ν_{max} /cm⁻¹ 3413 (N-H), 2982 (C-H), 1707 (C=O), 1632 (C=N); ¹H NMR (400 MHz, CDCl₃) δ _H 8.30 (s, 1H, Ar-H), 7.97 (s, 1H, Ar-H), 6.24 (s, 2H, adenosine-NH₂), 6.03 (d, *J*= 3.1, 1H, 1'-H), 5.69-5.65 (m, 1H, NHBoc), 5.40 (dd, *J*= 6.9, 1.4, 1H, 2'-H), 5.11 (dd, *J*= 6.0, 2.0, 1H, 3'-H), 4.47-4.43 (m, 1H, 4'-H), 3.27-3.22 (m, 2H, CH₂NHBoc), 3.30-3.26 (m, 2H, CH₂NH), 2.79 (dd, *J*= 12.0, 6.0, 1H, 5'-H_aH_b), 2.70 (dd, *J*= 12.0, 6.0, 1H, 5'-H_aH_b), 1.75-1.70 (m, 2H, CH₂CH₂CH₂), 1.55 (s, 3H, CH₃), 1.37 (s, 9H, C(CH₃)₃), 1.20 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ _C 171.0 (C), 156.1 (C), 152.1 (CH), 149.3 (C), 139.8 (CH), 120.3 (C), 114.5 (C), 90.8 (CH), 85.3 (CH), 83.2 (CH), 82.2 (CH), 78.8 (C), 47.8 (CH₂), 43.4 (CH₂), 39.2 (CH₂), 29.7 (CH₂), 28.4 (CH₃), 27.3 (CH₃), 25.4 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₁H₃₄N₇O₅⁺ 464.2616, found 464.2626.

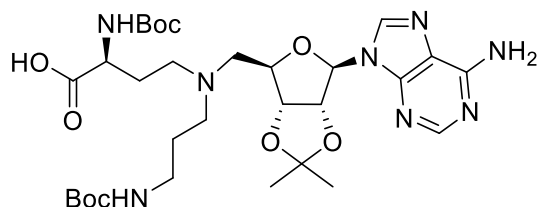
Preparation of benzyl(*S*)-4-((((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)(3-((*tert*-butoxycarbonyl)amino)propyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **44**



To a solution of secondary amine **43** (200 mg, 0.43 mmol) and aldehyde **36** (140 mg, 0.47 mmol) in 1,2-dichloroethane (9 mL) were mixed thoroughly for 15 min at room temperature. $\text{NaBH}(\text{OAc})_3$ (120 mg, 0.60 mmol) was added in one portion and stirred for 16 hours. The solution was quenched with saturated aqueous Na_2CO_3 (5 mL) and extracted with CH_2Cl_2 (2×10 mL), dried over Na_2SO_4 , filtered, and concentrated to give the product as white foam (0.24 g, 75 %): m.p. 84-91 °C; $[\alpha]_D^{25} -2.63$ (c 2.50 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (NH), 3009 (C-H, aromatic), 1706 and 1630 (C=O), 1161 (C-O); ^1H NMR (400 MHz, CD_3CN) δ_{H} 8.24 (s, 1H, adenosine-H), 8.03 (s, 1H, adenosine-H), 7.38-7.30 (m, 5H, *Ph*), 6.14 (br. s, 2H, NH_2), 6.10 (d, $J = 4.0$, 1H, 1'-H), 5.40 (dd, $J = 6.0, 1.3$, 1H, 2'-H), 5.17 (s, 2H, CH_2Ph), 5.04 (dd, $J = 6.0, 2.0$, 1H, 3'-H), 4.29-4.25 (m, 2H, 4'-H and CHNHOBoc), 3.00-2.96 (m, 2H, CH_2NBoc), 2.75-2.70 (m, 2H, CH_2CH), 2.56 (dd, $J = 13.1, 6.0$, 1H, 5'- $H_{\text{a}}H_{\text{b}}$), 2.41 (dd, $J = 13.1, 6.0$, 1H, 5'- $H_{\text{a}}H_{\text{b}}$), 2.37-2.28 (m, 4H, $2 \times \text{CH}_2\text{N}$), 1.67-1.63 (m, 2H, $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.58 (s, 3H, CH_3), 1.44 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.38 (s, 3H, CH_3); ^{13}C NMR (100 MHz, CD_3CN) δ_{C} 171.5 (C), 156.8 (C), 152.8 (CH), 149.1 (C), 140.4 (CH), 138 (C), 128.4 (CH), 127.4 (CH), 127.1 (CH), 121.3 (C), 119.9 (C), 11.8 (C), 113 (C) 90.4 (CH), 85.3 (CH), 83.9 (CH), 83.5 (CH), 79.8, (C), 78.7 (C), 66.41 (CH_2), 55.9 (CH_2), 52.3 (CH), 51.6 (CH_2), 50.5 (CH_2), 38.6 (CH_2), 28.5

(CH₃), 27.6 (CH₂), 26.4 (CH₂), 26.4 (CH₃), 25.3 (CH₃); HRMS (ESI+) *m/z*: (M⁺H⁺) calcd C₃₇H₅₄N₈O₉⁺ 755.4087, found 755.4088.

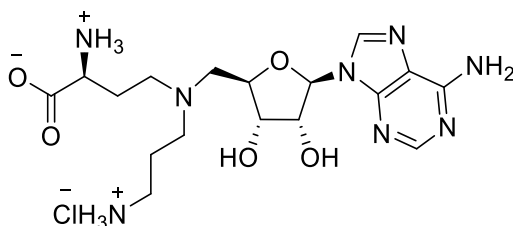
Preparation of (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-((tert-butoxycarbonyl)amino)propyl)amino)-2-((tert-butoxycarbonyl)amino)butanoic acid **45**



To a suspension of tertiary amine **44** (50 mg, 0.06 mmol) in methanol (1 mL) were added aqueous sodium hydroxide (2.0 M, 1 mL) slowly at room temperature. After stirring for 16 hours, the reaction was extracted with CHCl₃ (4 × 3 mL) and the aqueous layer was acidified using 1.0 M HCl until the pH adjusted to 5. The mixture reaction was extracted with CHCl₃ (4 × 2 mL). The organic layer was combined and dried over MgSO₄, filtered, concentrated *in vacuo* to give the product as pale yellow oil (0.03 mg, 68 %): m.p. 84-91 °C; [α]_D²⁵ + 23.6 (c 2.50 in CHCl₃); ν_{max} /cm⁻¹ 3413 (NH), 3009 (C-H, aromatic), 1706 and 1630 (C=O), 1161 (C-O); ¹H NMR (400 MHz, CD₃CN) δ_{H} 8.26 (s, 1H, adenosine-*H*), 8.10 (s, 1H, adenosine-*H*), 6.43 (br. s, 2H, NH₂), 6.19 (d, *J* = 2.0, 1H, 1'-*H*), 5.88 (m, 1H, NHBoc), 5.44 (dd, *J* = 6.0, 2.0, 1H, 2'-*H*), 5.11 (dd, *J* = 6.0, 2.0, 1H, 3'-*H*), 4.53 (m, 1H, 4'-*H*), 3.91-3.88 (m, 1H, CHNHBoc) 3.15 (m, 2H, CH₂NHBoc), 2.98 (dd, *J* = 14.6, 7.0, 1H, 5'-*H*_aH_b), 2.88 (dd, *J* = 14.6, 7.0, 1H, 5'-*H*_aH_b), 2.65-2.60 (m, 4H, 2 × CH₂N), 1.85 (ddd, 8.5, 7.0, 1.9, 1H, CH_aCH_b), 2.25 (ddd, *J* = 8.5, 7.0, 1.9,

^1H , CH_aCH_b), 2.01 (m, 2H, $\text{CH}_2\text{CH}_2\text{CH}_2$) 1.62 (s, 3H, CH_3), 1.49 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.42 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.35 (s, 3H, CH_3); ^{13}C NMR (100 MHz, CD_3CN) δ_{C} 171.5 (C), 156.8 (C), 152.8 (CH), 149.1 (C), 140.4 (CH), 138 (C), 127.4 (CH), 121.3 (C), 119.9 (C), 11.8 (C), 113 (C), 90.4 (CH), 85.3 (CH), 83.9 (CH), 83.5 (CH), 78.7 (C), 66.41 (CH_2), 55.9 (CH_2), 52.3 (CH), 51.6 (CH_2), 50.5 (CH_2), 38.6 (CH_2), 28.5 (CH_3), 27.6 (CH_2), 26.4 (CH_3), 25.3 (CH_3); HRMS (ESI+) m/z : (M^+H^+) calcd $\text{C}_{30}\text{H}_{48}\text{N}_8\text{O}_9^+$ 665.3617, found 665.3635.

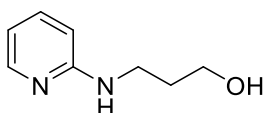
Preparation of (S)-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-ammoniopropyl)amino)-2-ammoniobutanoate chloride **46**



For deprotection of the bisubstrate inhibitor, protected tertiary amine **45** (40 mg, 0.06 mmol) was dissolved in neat TFA (5 mL) and stirred for 1 hour at room temperature. Water (0.5 mL) was then added and continually left to stirring for 4 hours. The reaction mixture was concentrated and dissolved in H_2O (5 mL). The mixture was then washed with EtOAc (2×5 mL) and the aqueous layer was concentrated. The resulting TFA salt was applied to Amberlite IRA 400 (Cl^-) ion exchange resin, eluting with H_2O to give the product as pale yellow solid (20 mg, 80 %): m.p. 84-91 $^\circ\text{C}$; $[\alpha]_{\text{D}}^{25} + 0.03$ (c 2.50 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3487 (N-H), 3008 (C-H, aromatic), 1602 (C=O), 1016 (C-O); ^1H NMR (400 MHz, D_2O) δ_{H} 8.28 (s, 1H, adenosine-H), 8.26 (s, 1H, adenosine-H), 6.06 (d, $J = 4.4$, 1H, 1'-H), 4.82

(app. t, $J = 4.4$, 1H, 2'-H), 4.41 (m, 2H, 3'-H and, 4'-H), 3.70-3.60 (m, 3H, CHNH₂ and CH₂NH₂), 3.45-3.39 (m, 4H, CH₂N), 2.96-2.902.21 (m, 4H, 5'-H_aH_b, 5'-H_aH_b and CH₂N), 2.96-2.90 (ddd, $J = 15.3, 9.0, 7.0$, 1H, CH_aCH_bCH), 2.10-2.02 (m, 3H, CH₂CH₂CH₂ and CH_aCH_bCH); ¹³C NMR (100 MHz, D₂O) δ_c 172.5 (C), 153.8 (C), 148.8 (CH), 148.1 (C), 142.4 (CH), 119.6 (C), 89.4 (CH), 78.3 (CH), 72.9 (CH), 71.5 (CH), 55.9 (CH₂), 53.3 (CH), 52.9 (CH₂), 52.1 (CH₂), 36.5 (CH₂), 24.6 (CH₂), 21.4 (CH₂); HRMS (ESI+) m/z : (M⁺H⁺) calcd C₁₇H₂₈N₈O₅⁺ 425.2255, found 425.2275.

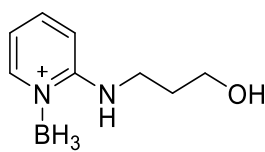
Preparation of 3-(pyridin-2-ylamino)propan-1-ol **48**⁷⁰



A solution of 2-chloropyridine (0.83 mL, 8.80 mmol) and solution of 3-aminopropanol (4.67 mL, 26.4 mmol) in a sealed tube was heated at 150 °C with continuous stirring overnight. Then, H₂O (10 mL) was added and the mixture was extracted with ethyl acetate (4 × 10 mL), washed with water (10 mL), dried over sodium sulfate and filtered through Celite and the solvent was removed *in vacuo*. The residue was purified by silica gel column chromatography using gradient elution with MeOH-CH₂Cl₂ (5/100 to 10/100, v/v) to give the product as pale yellow oil (890 mg, 67 %): $\nu_{\max}/\text{cm}^{-1}$ 3625 (O-H), 3458 (N-H), 3004 (C-H, aromatic); ¹H NMR (400 MHz, CDCl₃) δ_H 7.96 (ddd, $J = 7.0, 2.0, 0.9$, 1H, Ar-H), 7.34 (ddd, $J = 8.5, 7.0, 2.5$, 1H, Ar-H), 6.47 (ddd, $J = 7.0, 2.0, 0.9$, 1H, Ar-H), 6.37 (d, $J = 8.5$, 1H, Ar-H), 5.25 (s, 1H, NH), 5.05 (s, 1H, OH), 3.65 (t, $J = 5.7$, 2H, CH₂), 3.43 (app. q, $J = 5.7$, 2H, CH₂), 1.76-1.72 (m, 2H, CH₂); ¹³C NMR (100

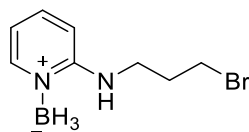
MHz, CDCl₃) δ_c 159.1 (C), 147.3 (CH), 137.4 (CH), 112.5 (CH), 108.1 (CH), 59.0 (CH₂), 38.5 (CH₂), 33.0 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd C₈H₁₃N₂O⁺ 153.1023, found 153.1030.

Preparation of 1-boranyl-2-((3-hydroxypropyl)amino)pyridin-1-ium 51⁷²



A solution of 3-(Pyridin-2-ylamino)propan-1-ol **48** (150 mg, 0.98 mmol) in anhydrous THF (5 mL) was cooled to 0 °C under argon. A solution of borane dimethylsulfide complex (2.00 M solution in THF, 0.18 mL, 1.96 mmol) was added over 10 min. After addition, the mixture was stirred at room temperature for 5 hours. Water (3 mL) was added then the mixture was extracted with ethyl acetate (3 × 5 mL). The organic extracts were concentrated under reduced pressure. The residue was subjected to flash column chromatography eluting with CH₂Cl₂/MeOH (98:2) to give product as a residual gum (115 mg, 72 %): $\nu_{\max}/\text{cm}^{-1}$ 3626 (O-H), 3010 (C-H, aromatic); ¹H NMR (400 MHz, CDCl₃) δ_H , 8.16 (d, J = 5.4, 1H, Ar-H), 7.61-7.57 (m, 1H, Ar-H), 6.67 (d, J = 8.8, 1H, Ar-H), 6.57-6.54 (m, 1H, Ar-H), 6.44 (s, br, 1H, NH), 5.31 (s, 1H, OH), 3.84 (app. t, J = 5.9, 2H, CH₂NH), 3.47 (app. q, J = 6.9, 2H, CH₂OH), 1.98-1.94 (m, 2H, CH₂CH₂CH₂), 2.3-2.0 (m, 3H, BH₃); ¹¹B NMR (128 MHz, CDCl₃) δ_B -2.0 (br. s, B, BH₃), 2.1 (br. s, 3H, BH₃); ¹³C NMR (100 MHz, CDCl₃) δ_c 154.9 (C), 146.3 (CH), 139.9 (CH), 111.30 (CH), 107.1 (CH), 59.99 (CH₂), 39.7 (CH₂), 31.30 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd, C₈H₁₆N₂OB⁺ requires 167.1350, found 167.1350.

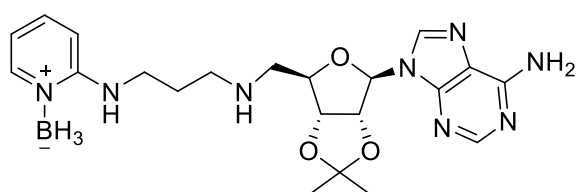
Preparation of 1-boranyl-2-((3-bromopropyl)amino)pyridine-1-ium **52**⁷²



A suspension of 3-(2-pyridinylamino)-1-propanol **48** (430 mg, 2.81 mmol) in dry CH₂Cl₂ (3 mL) was cooled to 0 °C under argon. A solution of thionyl bromide was added over 10 minutes, then the resultant mixture was stirred at room temperature for two hours under argon and concentrated *in vacuo*. The residue was dissolved in CH₂Cl₂ (3 mL), and the suspension cooled to -70 °C. Diisopropylethylamine (0.52 mL, 3.09 mmol) was added slowly while maintaining the reaction at -65 °C. After adding a solution of borane dimethylsulfide complex (0.33 mL, 3.09 mmol) was added slowly over 10 minutes. The reaction mixture was warmed to -10 °C, a saturated aqueous sodium hydrogen carbonate (3 mL) and H₂O (3 mL) were added. The mixture was then extracted with CH₂Cl₂ (2 × 4 mL). The organic layers were combined and concentrated *in vacuo*. The residue was dissolved in methanol (7 mL) and left stirring for 15 minutes and concentrated to remove any CH₂Cl₂. Addition of further methanol to the residue was carried immediately. Then, the precipitate formed was collected to provide as a white solid (450 mg, 76 %): $\nu_{\text{max}}/\text{cm}^{-1}$ 3458 (N-H), 3004 (C-H, aromatic); ¹H NMR (500 MHz, CDCl₃) δ_{H} , 8.24 (d, *J*= 5.7, 1H, Ar-*H*), 7.87-7.83 (m, 1H, Ar-*H*), 6.96 (d, *J*= 8.8, 1H, Ar-*H*), 6.80-6.76 (m, 1H, Ar-*H*), 6.48 (s, br, 1H, NH), 3.69-3.64 (m, 2H, CH₂Br), 2.33-2.30 (m, 2H, CH₂NH), 2.09-2.05 (m, 2H, CH₂CH₂CH₂), 2.33-2.07 (m, br, 3H, BH₃); ¹¹B NMR (128 MHz, CDCl₃) δ_{B} -2.0 (br. s, B, BH₃), 2.1 (br. s, 3H, BH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 146.1 (C), 140.5 (CH), 117.4 (CH), 111.8 (CH), 107.8 (CH), 40.6 (CH₂), 31.5 (CH₂), 31.2 (CH₂); HRMS (ESI+) *m/z*: (M-H⁺)

calcd, C₈H₁₃N₂¹¹B⁷⁹Br⁺ requires 227.0350 found 227.035 and C₈H₁₃N₂¹¹B⁸¹Br⁺ 229.0335 found 229.0338.

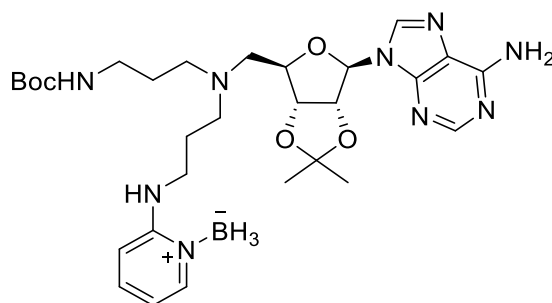
Preparation of 2-((3-(((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)amino)propyl)amino)-1-boraneypyridin-1-ium **53⁷²**



A solution of amine **33** (0.50 g, 1.94 mmol) was dissolved in acetonitrile (30 mL) and a solution of 3-(2-pyridinylamino)-1-propyl hydrobromide **52** (0.44 g, 1.62 mmol) and tetrabutylammonium iodide (0.02 g, 0.08 mmol) were added in the presence of potassium carbonate (0.89 g, 6.48 mmol) at room temperature under argon. The reaction mixture was heated to 54 °C for 18 hours and then cooled to room temperature. Crude product was concentrated under reduce of pressure by removal of solvent *in vacuo*. The residue was dissolved in *tert*-butyl methyl ether (TBME) (20 mL) with continuous mixing for 30 minutes. Then, water (20 mL) was added and left stirring for 30 minutes. The mixture was again extracted with (TBME) (4× 30 mL), dried over sodium sulfate, concentrated, and the residue was subjected to flash column chromatography using MeOH/CH₂Cl₂ (4/100, v/v) as eluent to provide the product as colourless oil (0.39 g, 52 %): [α]_D²² -0.28 (c 1.0 in CHCl₃); ν_{max} /cm⁻¹ 3411 (N-H), 1508 (C=C, aromatic), 1080 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.25 (s, 1H, adenosine-*H*), 8.16 (dd, *J*= 6.7, 1.2, 1H, Ar-*H*), 7.96 (s, 1H, adenosine-*H*), 7.59-7.54 (m, 1H, Ar-*H*), 6.68 (s, br, 1H, NH), 6.63 (d, *J*=

6.7, 1H, Ar-*H*), 6.57-6.53 (m, 1H, Ar-*H*), 6.03 (d, *J*= 3.4, 1H, 1'-*H*), 5.71 (s, br, 2H, NH₂) 5.51 (dd, *J*= 6.6, 3.4, 1H, 2'-*H*), 5.16 (dd, *J*= 6.6, 3.4, 1H, 3'-*H*), 4.43-4.40 (m, 1H, 4'-*H*), 3.40-3.36 (m, 2H, CH₂NH) 2.96 (m, 2H, CH₂NH-adenosine), 2.81 (dd, *J*= 12.1, 6.0, 1H, 5'-H_aH_b), 2.72 (dd, *J*=12.1, 6.0, 1H, 5'-H_aH_b), 1.83-1.79 (m, 2H, CH₂CH₂CH₂), 1.62 (s, 3H, CH₃), 1.39 (s, 3H, CH₃); ¹¹B NMR (128 MHz, CDCl₃) δ_B-2.0 (br. s, B, BH₃), 2.1 (br. s, 3H, BH₃); ¹³C NMR (100 MHz, CDCl₃) δ_C 156.0 (C), 154.9 (C), 152.5 (CH), 148.1 (C), 146.2 (CH), 140.6 (CH), 139.8 (CH), 119.3 (C), 114.2 (C), 110.9 (CH), 107.1 (CH), 90.4 (CH), 85.1 (CH), 83.1 (CH), 82.4 (CH), 50.9 (CH₂), 34.5 (CH₂), 29.7 (CH₂), 28.3 (CH₂), 26.0 (CH₃), 24.1 (CH₃); HRMS (ESI+) *m/z*: (M⁺H⁺) calcd C₂₁H₃₂BN₈O₃⁺ 455.2685, found 455.2687. The proton of borane is not visible in the proton NMR, but the borane is present in the ¹¹B spectra and mass as described in this data.

Preparation of 2-((3-(((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)(3-((*tert*-butoxycarbonyl)amino)propyl)amino)propyl)amino)-1-borane-1-ium 54

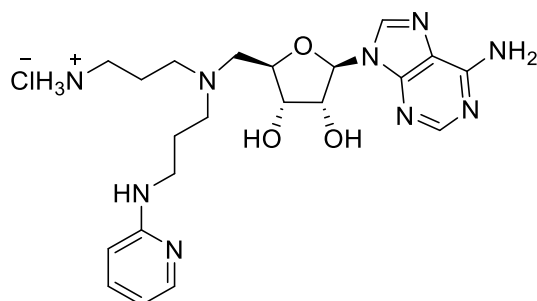


NaBH(OAc)₃ (192 mg, 0.91 mmol) was added slowly to a suspension of secondary amine **53** (200 mg, 0.43 mmol) and aldehyde **42** (140 mg, 0.86 mmol) in 1,2-dichloroethane (10 mL). The reaction mixture was stirred for 16 hours at room

temperature. After that the reaction was quenched with saturated aqueous Na_2CO_3 (10 mL), extracted with CH_2Cl_2 (2 × 20 mL), the combined organic layers were dried over Na_2SO_4 , filtered, concentrated *in vacuo* and the residue purified by silica gel chromatography using $\text{MeOH}/\text{CH}_2\text{Cl}_2$ (4/100, v/v) as eluent to provide the product as white foam (160 mg, 61 %): $[\alpha]^{23}_{\text{D}} + 6.79$ (c 1.0 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3412 (N-H), 3095 (C-H, aromatic), 1699 (C=O) 1507 (C=C, aromatic), 1035 (C-O); ^1H NMR (400 MHz, CD_3CN) δ_{H} 8.25 (s, 1H, adenosine-H), 8.06 (s, 1H, adenosine-H), 7.96 (dd, $J = 5.1, 1.0$, 1H, Ar-H), 7.38-7.35 (m, 1H, Ar-H), 6.50 (ddd, $J = 6.0, 5.1, 0.9$, 1H, Ar-H), 6.43 (d, $J = 8.5$, 1H, Ar-H), 6.20 (br. s, 2H, NH_2) 6.15 (d, $J = 1.9$, 1H, 1'-H), 5.56-5.50 (s, br, 2H, 2 × NH), 5.51 (dd, $J = 6.3, 1.9$, 1H, 2'-H), 5.03 (dd, $J = 6.3, 2.9$, 1H, 3'-H), 4.40-4.36 (m, 1H, 4'-H), 3.20 (app. t, $J = 6.4$, 2H, CH_2NHBoc), 3.05-3.00 (m, 2H, CH_2NH), 2.75 (dd, $J = 13.5, 6.8$, 1H, 5'- H_aH_b), 2.73 (dd, $J = 13.5, 8.0$, 1H, 5'- H_aH_b), 2.40 (br. m, 7H, 2 × CH_2N and BH_3), 1.63-1.50 (m, 4H, $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.59 (s, 3H, CH_3), 1.43 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.37 (s, 3H, CH_3); ^{11}B NMR (128 MHz, CDCl_3) δ_{B} -2.0 (br. s, B, BH_3), 2.1 (br. s, 3H, BH_3); ^{13}C NMR (100 MHz, CD_3CN) δ_{C} 161.5 (C), 156.0 (C), 152.8 (CH), 149.1 (C), 140.4 (CH), 137.2 (CH), 128.2 (C), 127.2 (C), 117.3 (CH), 114 (C), 113.7 (CH), 111.1 (CH), 90.3 (CH), 85.4 (CH), 83.3 (CH), 83.9 (CH), 80.2 (C), 55.7 (CH_2), 51.3 (CH_2), 51.1 (CH_2), 39.6 (CH_2), 33.8 (CH_2), 27.6 (CH_3), 26.3 (CH_3), 25.9 (CH_2), 24.1 (CH_2), 23.3 (CH_3); HRMS (ESI+) m/z : (M^+H^+) calcd $\text{C}_{29}\text{H}_{44}\text{N}_9\text{O}_5^+$ 598.3460, found 598.3446.

Borane observed by NMR spectroscopy, but molecular ion was not observed by electrospray mass spectrometry.

Preparation of 3-((((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-(pyridin-2-ylamino)propyl)amino)propan-1-aminium chloride **55**

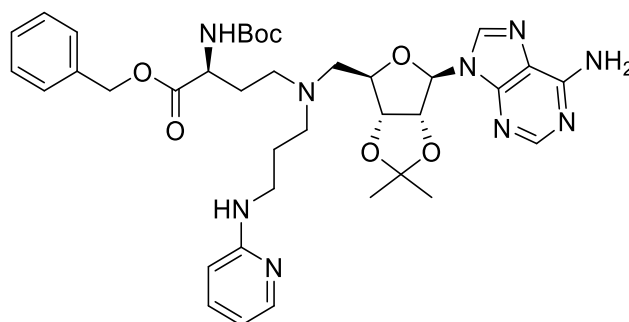


For deprotection, neat TFA (5 mL) was added to a suspension of protection tertiary amine **54** (40 mg, 0.06 mmol) and stirred for 1 hour at room temperature. After stirring for 1 hour, water (0.5 mL) was added and left with continual stirring for 4 hours at room temperature. The reaction mixture was concentrated and dissolved in H₂O (33 mL). Crude product was washed with EtOAc (2 × 3 mL) and the aqueous layer was concentrated under reduced pressure. The concentrated was re-dissolved in H₂O (5 mL) and washed with EtOAc (2 × 5 mL), and the aqueous layer was again concentrated. The resulting TFA salt was applied to Amberlite IRA 400 (Cl⁻) ion exchange resin, eluting with H₂O to give the product as pale yellow solid (20 mg, 80 %): $[\alpha]^{23}_{\text{D}} + 7.40$ (c 1.0 in H₂O); $\nu_{\text{max}}/\text{cm}^{-1}$, 3475 (N-H), 3009 (C-H, aromatic), 1465 (C=C, aromatic), 1016 (C-O); ¹H NMR (400 MHz, D₂O) δ_{H} 8.40 (s, 1H, adenosine-H), 8.35 (s, 1H, adenosine-H), 7.74-7.70 (m, 1H, Ar-H), 7.84 (d, *J* = 6.2, 1H, Ar-H), 6.62 (d, *J* = 6.2, 1H, Ar-H), 6.74-6.70 (m, 1H, Ar-H), 6.05 (d, *J* = 3.3, 1H, 1'-H), 4.61 (dd, *J* = 7.20, 3.3, 1H, 2'-H), 4.44 (m, 2H, NH and 1H, 3'-H), 3.82-3.78 (m, 1H, 4'-H), 3.66-3.60 (m, 2H, CH₂NH), 3.38-3.30 (m, 4H, 2x CH₂N), 3.03-2.99 (m, 2H, 5'-CH_aCH_b and 5'-CH_aCH_b), 2.11 (m, 2H, CH₂CH₂CH₂), 2.03 (m, 2H, CH₂CH₂CH₂); ¹³C NMR (100 MHz, D₂O) δ_{C} , 152.4 (C), 149.9 (C), 149.0 (CH) 148.1 (C), 147.1 (CH), 143.7 (CH), 143.2 (CH), 134.1 (CH), 119.7

(C), 112 (CH), 90.3 (CH), 84.4 (CH), 81.3 (CH), 71.9 (CH), 51.7 (CH₂), 38.3 (CH₂), 38.1 (CH₂), 35.6 (CH₂), 35.6 (CH₂), 35.1 (CH₂), 21.7 (CH₂); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₁H₃₁N₉O₃⁺ 458.2623, found 458.2623.

Preparation of benzyl (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-(pyridin-2-ylamino)propyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate

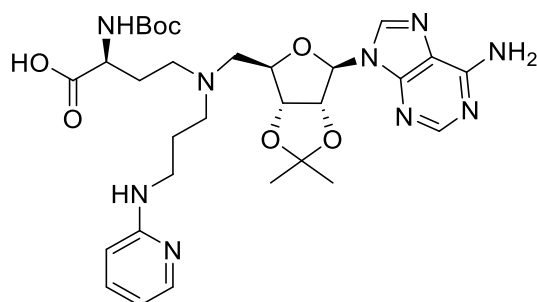
56



NaBH(OAc)₃ (180 mg, 0.87 mmol) was added in one portion to a suspension of secondary amine **53** (200 mg, 0.43 mmol) and aldehyde **36** (260 mg, 0.87 mmol) in 1,2-dichloroethane (10 mL) at room temperature. After stirring 16 hours, the mixture was quenched with saturated aqueous Na₂CO₃ (10 mL) and then extracted with dichloromethane (4 × 20 mL). The organic layers were dried over Na₂SO₄, filtered, concentrated *in vacuo* and the resulting residue was subjected to flash column chromatography (silica gel, MeOH/CH₂Cl₂ 4/100, v/v) as eluent to give the product as white foam (180 mg, 56 %): m.p. 94-97 °C; [α]_D²³ + 25.8 (c 1.0 in CHCl₃); ν_{max} /cm⁻¹ 3400 (N-H), 3094 (C-H, aromatic), 1630 (C=O) 1240 (C-N), 1035 (C-O); ¹H NMR (400 MHz, CD₃OD) δ_{H} 8.23 (s, 1H, adenosine-H), 8.19 (s, 1H, adenosine-H), 7.89 (dd, *J* = 5.3, 1.0, 1H, Ar-H), 7.41-7.37 (m, 1H, Ar-H) 7.27-

7.23 (m, 5H, *Ph*), 6.53 (ddd, $J = 6.2, 5.3, 0.8$, 1H, Ar-*H*), 6.48 (d, $J = 8.5$, 1H, Ar-*H*) 6.17 (d, $J = 2.4$, 1H, 1'-*H*), 5.43 (dd, $J = 6.5, 2.4$, 1H, 2'-*H*), 5.20 (s, 2H, CH₂Ph), 5.06 (dd, $J = 6.5, 3.2$, 1H, 3'-*H*), 4.36-4.32 (m, 1H, 4'-*H*), 4.33-4.30 (m, 1H, CHNH*Boc*), 3.36 (s, 1H, NH), 3.33 (app. t, $J = 6.7$, 2H, CH₂NH), 2.80 (dd, $J = 13.5, 6.8$, 1H, 5'-CH_aCH_b), 2.65-6.58 (m, 3H, CH₂N and 5'-CH_aCH_b), 2.47-2.42 (m, 2H, CH₂N), 2.05-2.01 (m, 1H, CH_aCH_bCH), 1.65-1.77 (m, 3H, CH₂CH₂CH₂ and CH_aCH_bCH), 1.62 (s, 3H, CH₃), 1.42 (s, 9H, C(CH₃)₃), 1.35 (s, 3H, CH₃); ¹³C NMR (100 MHz, CD₃OD) δ_c 171 (C), 161.5 (C), 156.0 (C), 155.1 (C), 154.1 (CH), 152.8 (CH), 149.1 (C), 140.4 (CH), 137.2 (CH), 132.1 (C), 128.2 (CH), 127.5 (CH), 127.0 (CH), 119.2 (C), 119.7 (CH), 114 (C), 113.1 (CH), 108 (CH), 90.3 (CH), 85.4 (CH), 83.3 (CH), 83.9 (CH), 80.2 (C), 55.7 (CH₂), 51.3 (CH₂), 51.1 (CH₂), 39.6 (CH₂), 33.8 (CH₂), 27.6 (CH₃), 26.3 (CH₃), 25.9 (CH₂), 24.1 (CH₂), 23.3 (CH₃); HRMS (ESI+) m/z : (M+H⁺) calcd C₃₇H₅₀N₉O₇⁺ 732.3828, found 732.3847. The evidence from mass and NMR spectrometry confirm that the borane-pyridinium complex was unstable and decomposed with this compound during the reaction.

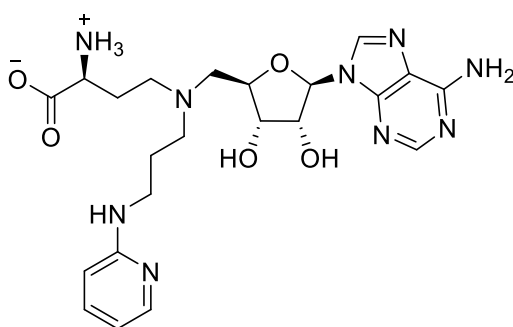
Preparation of (S)-4-((((3aR,4R,6R,6aR)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-(pyridin-2-ylamino)propyl)amino)-2-((tert-butoxycarbonyl)amino)butanoic acid
57



Aqueous sodium hydroxide (2.0 M, 2 mL) was added slowly to a suspension of tertiary amine **56** (40 mg, 0.054 mmol) in methanol at room temperature. After stirring for 16 hours, the reaction was extracted with CHCl_3 (4 $\times$ 3 mL) and the aqueous layer was acidified using 1.0 M HCl until pH 5.0. The acidified product was then extracted with CHCl_3 (4 $\times$ 2 mL). The organic layer was dried over MgSO_4 , filtered, concentrated *in vacuo* to give the product as pale yellow oil (20 mg, 68 %): m.p. 99-102 °C; $[\alpha]^{23}_{\text{D}} + 26.4$ (c 1.0 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (N-H), 2983 (C-H, aromatic), 1698 (C=O) 1240 (C-N), 1078 (C-O); ^1H NMR (400 MHz, CDCl_3) δ_{H} 8.30 (s, 1H, adenosine-H), 8.00 (s, 1H, adenosine-H), 7.84 (d, J = 5.12, 1H, Ar-H), 7.48-7.44 (m, 1H, Ar-H) 6.53-6.50 (m, 1H, Ar-H), 6.48 (d, J =8.6, 1H, Ar-H) 6.36 (br. s, 2H, NH_2) 6.08 (d, J = 1.9, 1H, 1'-H), 5.93 (d, J = 5.85, 1H, NHBoc), 5.45 (dd, J = 5.8, 1.9, 1H, 2'-H), 5.03 (dd, J = 5.90, 3.8, 1H, 3'-H), 4.50-4.46 (m, 1H, 4'-H), 4.14-4.10 (m, 1H, CHNHBoc), 3.22-3.19 (m, 2H, CH_2NH), 2.95 (d, J = 5.5, 1H, NH), 2.80 (dd, J = 12.8, 6.7, 1H, 5'- CH_aCH_b), 2.65-6.58 (m, 3H, CH_2N and 5'- CH_aCH_b), 2.47-2.42 (m, 2H, CH_2N), 1.99-1.95 (m, 2H, $\text{CH}_a\text{CH}_b\text{CH}$ and $\text{CH}_a\text{CH}_b\text{CH}$), 1.77-1.73 (m, 2H, $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.60 (s, 3H, CH_3), 1.44 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.38 (s, 3H, CH_3); ^{13}C NMR (100 MHz, CDCl_3) δ_{C} 174 (C), 157.5 (C),

156.0 (C), 155.1 (C), 152.8 (CH), 147.1 (C), 137.2 (CH), 132.1 (CH), 128.2 (CH), 119.2 (C), 114 (C), 111.1 (CH), 108 (CH), 90.3 (CH), 84.9 (CH), 84.3 (CH), 83.1 (CH), 80.0 (C), 55.7 (CH₂), 53.2 (CH), 51.3 (CH₂), 51.1 (CH₂), 40.1 (CH₂), 38.6 (CH₂), 29.9 (CH₂), 28.6 (CH₃), 27.3 (CH₃), 25.3 (CH₃), 24.1 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd C₃₀H₄₃N₉O₇⁺ 642.3358, found 642.3365.

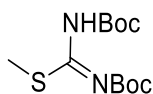
Preparation of (S)-4-((((2R,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-(pyridin-2-ylamino)propyl)amino)-2-ammoniobutanoate **58**



For deprotection of the bisubstrate inhibitor, protected tertiary amine **57** (116 mg, 0.15 mmol) was dissolved in neat TFA (5 mL) and stirred for 1 hour at room temperature. H₂O (0.5 mL) was then added and stirring continued for 4 hours. The reaction mixture was concentrated and dissolved in H₂O (5 mL), washed with EtOAc (2 × 5 mL) and the aqueous layer was concentrated. The resulting was subjected to Amberlite IRA 400 (Cl⁻) ion exchange resin, eluting with H₂O to give the product as pale-yellow solid (25 mg, 85 %): m.p. 94-97 °C; [α]_D²³ + 20.3 (c 1.0 in H₂O); $\nu_{\max}/\text{cm}^{-1}$ 3441 (N-H), 3007 (C-H, aromatic), 1623 (C=O), 1240 (C-N), 1015 (C-O); ¹H NMR (400 MHz, D₂O) δ_{H} 8.39 (s, 1H, adenosine-H), 8.34 (s, 1H, adenosine-H), 7.74-7.70 (m, 1H, Ar-H), 7.84 (d, J = 6.2, 1H, Ar-H), 6.75 (d,

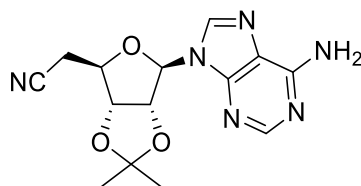
$J = 6.75$, 1H, Ar- H), 6.74-6.70 (m, 1H, Ar- H), 6.04 (d, $J = 3.5$, 1H, 1'- H), 4.61 (m, 2H, 2'- H and 1H, 3'- H), 4.04 (m, 1H, 4'- H), 3.80 (m, 1H, $CHNH_2$), 3.69-3.60 (m, 2H, CH_2NH), 3.38 (m, 6H, 5'- CH_aCH_b and 5'- CH_aCH_b and $2 \times CH_2N$), 2.38 (ddd, 14.4, 10.2, 8.6, 1H, CH_aCH_bCH), 2.25 (ddd, $J = 14.4$, 10.6, 5.0, 1H, CH_aCH_bCH), 2.05 (m, 2H, $CH_2CH_2CH_2$); ^{13}C NMR (100 MHz, D_2O) δ_c 185.1 (C), 170.4 (C), 155 (C), 152.0 (CH), 149.1 (C), 147.1 (CH), 144.8 (CH), 143.7 (CH), 143.2 (CH), 134.1 (CH), 119.2 (C), 90.3 (CH), 78.4 (CH), 73.3 (CH), 71.9 (CH), 54.7 (CH_2), 52.3 (CH_2), 52.1 (CH), 51.1 (CH_2), 38.6 (CH_2), 29.2 (CH_2), 24.1 (CH_2); HRMS (ESI+) m/z : ($M+H^+$) calcd $C_{22}H_{31}N_9O_5^+$ 502.2520, found 502.2521.

Preparation of *N,N'*-bis-(*tert*-Butoxycarbonyl)-*S*-methylthiourea **60**⁷³



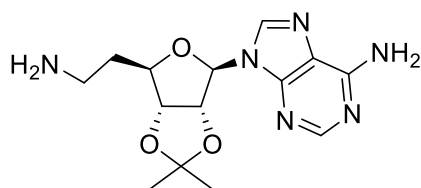
Saturated aqueous Na_2CO_3 (50 mL) was added to solution of *S*-methylisothiourea sulfate (2.00 g, 7.20 mmol) in CH_2Cl_2 (50 mL). Solution of di-*tert* butyl dicarbonate (6.3 g, 28.8 mmol) was then added slowly and stirred vigorously for 16 hours. The reaction mixture was extracted with CH_2Cl_2 (2×20 mL), and the aqueous layer was separated. The organic layer was combined and dried over Na_2SO_4 , filtered, evaporated and subjected to column chromatography using EtOAc/Petrol (1/3, v/v) to give the product as white solid (3.10 g, 74 % yield), m.p. 118-120 °C [lit. 122-123 °C];⁷² ν_{max}/cm^{-1} , 3170 (N-H), 2985 (C-H), 1811 and 1743 (C=O), 1648 (C=N); ^1H NMR (400 MHz, $CDCl_3$) δ_H 2.40 (s, 3H, CH_3), 1.53 (s, 9H, $C(CH_3)_3$), 1.51 (s, 9H, $C(CH_3)_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ_c 171.3 (C), 161.0 (C), 150.6 (C), 84.0 (C), 80.9 (C), 27.1 ($CH_3(CH_3)_6$), 14.4 (CH_3); HRMS (ESI+) m/z : ($M+Na^+$) calcd $C_{12}H_{22}N_2NaO_4S^+$ 313.1201, found 313.1197.

Preparation of 2-((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3] dioxol-4-yl)acetonitrile **61⁶²**



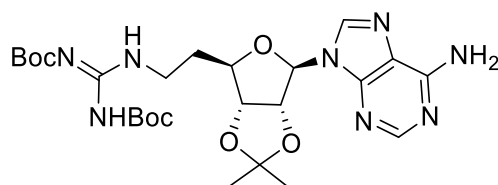
Acetone cyanohydrin (0.70 g, 8.15 mmol) was added within 5 min to a well stirred suspension of **30** (1 g, 3.25 mmol), PPh₃ (2.13 g, 8.12 mmol) and DEAD (1.39 g, 2.46 mmol) in THF (20 mL) at 0 °C. The solution was stirred at 0 °C for 5 min and further at 20 °C for 16 hours. The reaction mixture was evaporation *in vacuo* and suspended to column chromatography (SiO₂: EtOAc. MeOH (96:4) to afforded compound **61** as a brown oil (0.70 g, 67 %): $\nu_{\text{max}}/\text{cm}^{-1}$, 3041 (C-H), 1632 (C=N); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.36 (s, 1H, adenosine-*H*), 7.91 (s, 1H, adenosine-*H*), 6.14 (d, *J*= 2.08, 1H, 1'-*H*), 5.83 (br. s, 2H, NH₂), 5.48 (dd, *J*= 5.9, 1.0, 1H, 2'-*H*), 5.13 (dd, *J*= 5.9, 0.9, 1H, 3'-*H*), 4.55-4.52 (m, 1H, 4'-*H*), 3.00 (dd, *J*= 12.9, 1.7, 1H, 5'-CH_aH_b), 2.93 (dd, *J*= 12.9, 1.9, 1H, 5'-CH_aH_b), 1.64 (s, 3H, CH₃), 1.41 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 155.9 (C), 153.0 (CH), 148.4 (C), 140.2 (CH), 120.0 (C), 116.0 (C), 114.0 (C), 91.2 (CH), 84.1 (CH), 83.0 (CH), 82.0 (CH), 67.0 (CH₂), 27.2 (CH₃), 25.2 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd. C₁₄H₁₇N₆O₃⁺ 317.1353, found 317.1345.

Preparation of 9-((3a*R*,4*R*,6*R*,6a*R*)-6-(2-aminoethyl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)-9*H*-purin-6-amine **62**⁶²



Glacial acetic acid (20 mL) and PtO₂ (200 mg) were added to the stirred suspension of **61** (0.70 g, 2.21 mmol) in THF (11 mL) at room temperature. The suspension was hydrogenated at (4 bar H₂) for 19 hours to produce primary amine. After reaction was completed, the catalyst was removed by filtration and the solvent evaporated *in vacuo*. The residue was then purified by column chromatography (SiO₂) (CH₂Cl₂, MeOH) to afford (0.40 g 56 %) as a colourless foam. m.p. 105-107 °C; $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (N-H), 3008 (C-H), 1631 (C=N), 1082 (C-O); ¹H NMR (400 MHz, CD₃OD) δ_{H} 8.28 (s, 1H, adenosine-H), 8.24 (s, 1H, adenosine-H), 6.16 (d, *J*= 2.0, 1H, 1'-H), 5.48 (dd, *J*= 6.0, 2.0, 1H, 2'-H), 4.94 (dd, *J*= 5.9, 0.9, 1H, 3'-H), 4.24-4.23 (m, 1H, 4'-H), 2.73 (dd, *J*= 12.0, 2.0, 1H, 5'-CH_aH_b), 2.70 (dd, *J*= 12.0, 2.0, 1H, 5'-CH_aH_b), 2.10 (m, 2H, CH₂), 1.61 (s, 3H, CH₃), 1.38 (s, 3H, CH₃); ¹³C NMR (100 MHz, CD₃OD) δ_{C} 155.7 (C), 152.6 (CH), 148.8 (C), 140.8 (CH), 119.5 (C), 114.3 (C), 89.5 (CH), 84.7 (CH), 83.5 (CH), 83.0 (CH), 37.8 (CH₂), 34.6 (CH₂), 26.4 (CH₃), 23.2 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd. C₁₄H₂₁N₆O₃⁺ 321.1700, found 321.1681. The literature does not show the melting point.⁶²

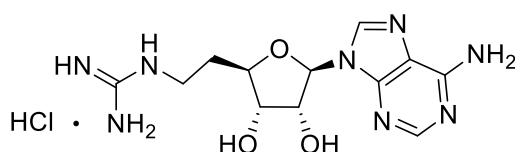
Preparation of 2-(2-((3a*R*,4*R*,6*R*,6a*R*)-6-(6-Amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)ethyl)-*N,N'*-di-Boc-guanidine **63⁶²**



NEt₃ (416 μ L, 3.0 mmol) was added slowly to a suspension of amine **62** (570 mg, 1.5 mmol) and a solution of *N,N'*-bis (*tert*-butoxycarbonyl) - *N''*-triflylguanidine **60** (392 mg, 1.0 mmol) in CH₂Cl₂ (5 mL). The mixture reaction was left stirred for 16 hours at room temperature and then diluted with CH₂Cl₂ (15 mL). The mixture was washed with saturated aqueous KHSO₄ (10 mL), saturated aqueous NaHCO₃ (10 mL) and brine (10 mL). The organic layer was separated and dried over Na₂SO₄, filtered, concentrated *in vacuo*. The residue was then subjected to column chromatography (SiO₂: EtOAc–MeOH (98: 2)) to produce compound **62** (422 mg, 75 %) as a white powder. m.p. 150-153 °C; ν_{max} /cm⁻¹ 3684 (N-H), 3033 (C-H), 1811 and 1743 (C=O), 1648 (C=N), 1234 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.36 (s, 1H, adenosine-*H*), 7.94 (s, 1H, adenosine-*H*), 6.07 (d, *J*= 6.3, 1H, 1'-*H*), 5.67 (br. s, 2H, NH₂), 5.49 (dd, *J*= 5.9, 3.0, 1H, 2'- *H*), 4.96 (dd, *J*= 5.9, 0.9, 1H, 3'-*H*), 4.27-4.26 (m, 1H, 4'-*H*), 3.67 (dd, *J*= 12.9, 1.7, 1H, 5'-CH_aH_b), 3.44 (dd, *J*= 12.9, 1.9, 1H, 5'-CH_aH_b), 2.07 (m, 2H, CH₂), 1.89 (br. s, 1H, NH), 1.62 (s, 9H, C(CH₃)₃), 1.50 (s, 9H, C(CH₃)₃), 1.61 (s, 3H, CH₃), 1.38 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 163.0 (C), 158.0 (C), 155.7 (C), 155.0 (C), 153.0 (C), 152.0 (CH), 149.0 (C), 140.0 (C), 124.7 (C), 114.0 (C), 90.0 (CH), 87.7 (CH), 85.7 (CH), 83.5 (CH), 82.0 (C), 71.3 (CH), 38.8 (CH₂), 33.6 (CH₂), 28.4 (CH₃), 27.2

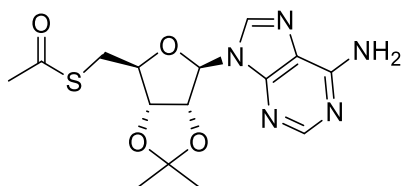
(CH₃), 26.4 (CH₃), 25.2 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd. C₂₅H₃₉N₈O₇⁺ 563.2953, found 563.2943. The literature does not show the melting point.⁶²

Preparation of 1-((2*R*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)ethyl) guanidine **64⁶²**



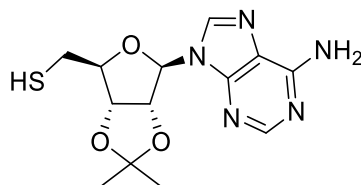
Compound **63** (56.5 mg, 0.1 mmol) was dissolved in neat TFA (5 mL) and stirred for 1 hour at room temperature. Water (0.5 mL) was then added to the suspension and stirring continued for 4 hours. The reaction mixture was concentrated *in vacuo* and the residue was then again diluted with water (5 mL). The aqueous layer was washed with EtOAc (2 × 5 mL). The aqueous layer was concentrated, and the resulting was subjected to Amberlite IRA 400 (Cl⁻) ion exchange resin, eluting with water to give the product **64** as a white powder (38.8 mg, 89 %). m.p. 162-164 °C; $\nu_{\text{max}}/\text{cm}^{-1}$ 3648 (N-H), 3008 (C-H), 1602 (C=N), 1242 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.32 (s, 1H, adenosine-*H*), 8.27 (s, 1H, adenosine-*H*), 6.03 (d, *J*= 4.7, 1H, 1'-*H*), 4.12 (m, 2H, 2'-*H* + 3'-*H*) 3.86-3.82 (m, 1H, 4'-*H*), 3.28 (m, 2H, CH₂), 2.34 (dd, *J*= 12.9, 1.7, 1H, 5'-CH_aH_b), 2.22 (dd, *J*= 12.9, 1.9, 1H, 5'-CH_aH_b); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 157.7 (C), 150.0 (C), 148.5 (C), 144.7 (C), 143.3 (C), 118.0 (C), 90.0 (CH), 85.7 (CH), 74.5 (CH), 74.0 (CH), 38.8 (CH₂), 33.6 (CH₂); HRMS (ESI+) *m/z*: (M+H⁺) calcd. C₁₂H₁₉N₈O₃⁺ 323.1600, found 323.1517. (The literature does not show the melting point).⁶²

Preparation of *S*-(((3a*S*,4*S*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)ethanethioate **65⁷⁹**



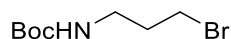
Diisopropylazodicarboxylate (5.11 mL, 26.0 mmol) was added dropwise over 5 minutes to a solution of PPh_3 (6.80 g, 26.0 mmol) in dry THF (40 mL) at 0 °C and stirred for 30 minutes under argon. A suspension of 2', 3'-*O*-isopropylideneadenosine **30** (4.00 g, 13.0 mmol) was added in one portion to the mixture reaction, and stirred for further 10 minutes at 0 °C. A solution of thioacetic acid (1.82 mL, 26.0 mmol) in dry THF (8 mL) was added dropwise over 5 minutes to the reaction and stirring continued for another 1 h at 0 °C. The solvent was evaporated to dryness *in vacuo*. The residue was subjected to the column chromatography using MeOH / CH_3Cl (4/100: v/v) to give as a yellow foam solid (4.45 g, 93 %): m.p. 90-92 °C; $[\alpha]^{25}_{\text{D}}$ -29.50 (c 1.0 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3441 (N-H), 2993 (C-H), 1694 (C=O) 1239 (C-N), 1093 (C-O); ^1H NMR (400 MHz, CDCl_3) δ_{H} 8.28 (s, 1H, adenosine-*H*), 7.90 (s, 1H, adenosine-*H*), 6.86 (s, 2H, adenosine- NH_2), 6.05 (d, $J = 2.1$, 1H, 1'-*H*), 5.49 (dd, $J = 6.3, 2.1$, 1H, 2'-*H*), 4.95 (dd, $J = 6.3, 3.1$, 1H, 3'-*H*), 4.29-4.24 (m, 1H, 4'-*H*), 3.27 (dd, $J = 13.8, 6.0$, 1H, 5'- H_aH_b), 3.16 (dd, $J = 13.8, 6.0$, 1H, 5'- H_aH_b), 2.28 (s, 3H, CH_3CO) 1.53 (s, 3H, CH_3), 1.33 (s, 3H, CH_3); ^{13}C NMR (100 MHz, CDCl_3) δ_{C} 194.5 (C), 156.1 (C), 153.0 (CH), 149.0 (C), 139.8 (CH), 120.1 (C), 114.3 (C), 90.8 (CH), 86.0 (CH), 84.15 (CH), 83.6 (CH), 31.24 (CH_2), 30.53 (CH_3), 26.02 (CH_3), 25.3 (CH_3); HRMS (ESI+) m/z : ($\text{M}+\text{H}^+$) calcd $\text{C}_{15}\text{H}_{20}\text{N}_5\text{O}_4\text{S}^+$ 366.1231, found 366.1218. (The literature does not show the $[\alpha]_{\text{D}}$ and melting point).⁷⁹

Preparation of ((3a*S*,4*S*,6*R*,6*R*))-6-(6-amino-9*H*-purine-9-yl)2,2-dimethyltetrahydrofuro{3, 4-*d*} {1, 3}dioxol-4-yl)methanethiol **66⁷⁹**



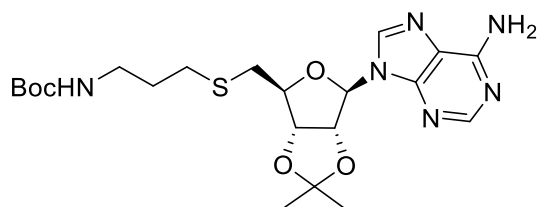
A solution of 5-acetylthio-5-deoxy-2',3'-O-isopropylideneadenosine **65** (1.00 g, 2.74 mmol) was dissolved with methanol in saturated with ammonia (15 mL) and stirred for 30 min at room temperature under argon. The reaction mixture was evaporated under reduce pressure to obtain the product as white foam (860 mg, 95 %). This product was then used directly in the next step without purification: $[\alpha]^{23}_{\text{D}} -0.42$ (c 1.0 in CHCl_3); $\nu_{\text{max}}/\text{cm}^{-1}$ 3414 (N-H), 2996 (C-H), 1239 (C-N), 1083 (C-O); ^1H NMR (400 MHz, CDCl_3) δ_{H} 8.23 (s, 1H, adenosine-*H*), 7.91 (s, 1H, adenosine-*H*), 6.86 (s, 2H, adenosine- NH_2), 6.50 (s, 1H, *SH*) 6.04 (d, $J = 1.8$, 1H, 1'-*H*), 5.52 (dd, $J = 6.3, 1.8$, 1H, 2'-*H*), 5.04 (dd, $J = 6.3, 2.8$, 1H, 3'-*H*), 4.45-4.40 (m, 1H, 4'-*H*), 3.00 (dd, $J = 13.9, 6.0$, 1H, 5'- H_aH_b), 2.91 (dd, $J = 13.9, 6.0$, 1H, 5'- H_aH_b), 1.57 (s, 3H, CH_3), 1.35 (s, 3H, CH_3); ^{13}C NMR (100 MHz, CDCl_3) δ_{C} 156.1 (C), 153.0 (CH), 149.0 (C), 140.0 (CH), 120.7 (C), 114.5 (C), 90.5 (CH), 86.1 (CH), 84.0 (CH), 83.0 (CH), 41.0 (CH_2), 25.3 (CH_3), 22.6 (CH_3); HRMS (ESI+) m/z : ($\text{M} + \text{H}^+$) calcd $\text{C}_{13}\text{H}_{18}\text{N}_5\text{O}_3\text{S}^+$ 324.1125, found 324.1119. (The literature does not show the $[\alpha]_{\text{D}}$).⁷⁹

Preparation of *tert*-butyl(3-bromopropyl)carbamate 69



Triethylamine (1.28 mL, 9.12 mmol) was added dropwise over one hour to a solution of 3-bromopropylamine (1.00 g, 4.56 mmol) and di-*tert*-butyl dicarbonate (1.04 mL, 4.56 mmol) in CH₂Cl₂ (3 mL) at 0 °C. The reaction was stirred at room temperature for 16 hours under argon. The mixture was diluted with CH₂Cl₂ (2 mL) and washed with 1.0 M HCl (2 × 2 mL) and water (2 × 2 mL). The organic layer was then dried over sodium sulfate, filtered, and evaporated to afford the product as a yellow oil (1.00 g, 92 %): $\nu_{\text{max}}/\text{cm}^{-1}$ 3458 (N-H), 3008 (C-H), 1711 (C=O), 1267 (C-N); ¹H NMR (400 MHz, CDCl₃) δ_{H} 4.88 (br. s, 1H, *NHBoc*), 3.38 (t, *J* = 6.6, 2H, CH₂Br), 3.22 (app. q, *J* = 6.3, 2H, CH₂NH), 2.02-1.97 (m, 2H, CH₂CH₂CH₂), 1.42 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 155.6 (C), 79.7 (C), 39.9 (CH₂), 32.7 (CH₂), 30.7 (CH₂), 28.1 (CH₃); HRMS (ESI+) *m/z*: (M+Na⁺) calcd C₈H₁₇NO₂Na⁷⁹Br 260.0262 found 260.0249, C₈H₁₇NO₂Na⁸¹Br 262.0242 found 262.0229.

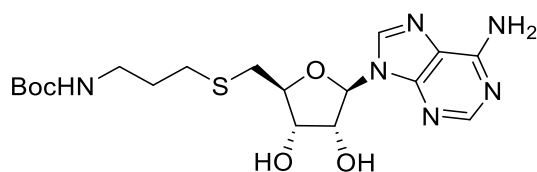
Preparation of *tert*-butyl(3-((((3*aS*,4*S*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)methyl)thio)propyl)carbamate **70**



A suspension of *tert*-butyl (3-bromopropyl) carbamate **69** (3.54 g, 14.9 mmol) in DMF (5 mL) was added slowly to a suspension of 5'-thio-5'-deoxy-2',3'-O-isopropylidene adenosine **66** (3.23 g, 9.96 mmol) in the presence of potassium carbonate (2.75 g, 19.9 mmol) with DMF (20 mL) at room temperature under argon. Sodium borohydride (0.26 g, 9.96 mmol) was added to the mixture reaction to reduce the disulfide bond of starting material. The mixture was then heated to 100 °C for 30 minutes and left cooled at room temperature. Water (20 mL) was added to the mixture and then extracted with ethyl acetate (3 × 40 mL). The organic layers were washed with water (20 mL) and brine (20 mL), dried over Na₂SO₄, filtered and evaporated to dryness, The residue was purified by silica gel chromatography using methanol/dichloromethane (2/100, v/v) as eluent to give the product as pale yellow oil (3.06 g, 63 %): $[\alpha]^{23}_{\text{D}}$ -13.0 (c 1.0 in CHCl₃); $\nu_{\text{max}}/\text{cm}^{-1}$ 3413 (N-H), 2984 (C-H), 1708 (C=O), 1241 (C-N), 1091 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.31 (s, 1H, Ar-*H*), 7.94 (s, 1H, Ar-*H*), 6.41 (s, 2H, adenosine-NH₂), 6.09 (d, *J*= 2.5, 1H, 1'-*H*), 5.52 (dd, *J*= 6.3, 2.5, 1H, 2'-*H*), 5.07 (dd, *J*= 6.3, 3.1, 1H, 3'-*H*), 4.37-4.32 (m, 1H, 4'-*H*), 3.15-3.10 (m, 2H, CH₂NHBoc), 2.79 (dd, *J*= 13.6, 6.0, 1H, 5'-*H_aH_b*), 2.73 (dd, *J*= 13.6, 6.0, 1H, 5'-*H_aH_b*), 2.52 (t, *J*= 7.2, 2H, CH₂S), 1.69-1.64 (m, 2H, CH₂CH₂CH₂), 1.59 (s, 3H, CH₃), 1.41 (s, 9H, C(CH₃)₃), 1.38 (s, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 157.2 (C), 155.9 (C),

153.1 (CH), 149.1 (C), 139.9 (CH), 120.3 (C), 114.5 (C), 90.8 (CH), 86.7 (CH), 88.0 (CH), 83.8 (CH), 77.4 (C), 53.4 (CH₂), 39.3 (CH₂), 34.3 (CH₂), 29.68 (CH₂), 29.5 (CH₃), 27.0 (CH₃), 25.3 (CH₃); HRMS (ESI+) m/z : (M+H⁺) calcd C₂₁H₃₂N₆O₅S⁺ 481.2228, found 481.2232.

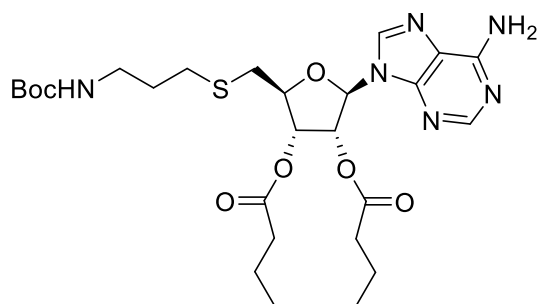
Preparation of *tert*-butyl(3-((((2*S*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)thio)propyl)carbamate **72**



Et₃N (0.06 mL, 0.45 mmol) was added dropwise to a solution of compound **71** (0.20 g, 0.45 mmol) in dry MeOH (5 mL) at room temperature and followed by slow addition of di-*tert*-butyl dicarbonate (0.10 mL, 0.45 mmol). The reaction was stirred at room temperature for 16 hours under argon. The mixture was diluted with CH₂Cl₂ (5 mL) and washed with water (5 × 2 mL). The organic layers were combined and then dried over sodium sulfate, filtered, and evaporated under reduce pressure. The residue was subjected to flash column chromatography using MeOH/CH₂Cl₂ (2/100, v/v) as eluent to give the product as pale yellow oil (0.16 g, 65 %): [α]_D²² + 22.7 (c 1.0 in CHCl₃); ν_{max} /cm⁻¹ 2094 (C-H), 1706 (C=O), 1240 (C-N), 1100 (C-O); ¹H NMR (400 MHz, CD₃CN) δ_{H} 8.24 (s, 1H, Ar-*H*), 8.08 (s, 1H, Ar-*H*), 6.10 (s, 2H, adenosine-NH₂), 5.94 (d, J = 5.0, 1H, 1'-*H*), 5.34 (br. s, 1H, NHBoc) 4.76 (app. t, J = 5.0, 1H, 2'-*H*), 4.34 (app. t, J = 5.3, 1H, 3'-*H*) 4.13-4.10 (m, 1H, 4'-*H*), 3.08-3.05 (m, 2H, CH₂NHBoc), 2.91 (dd, J =14.0, 5.7, 2H, 5'-H_aH_b), 2.85 (dd, J = 14.0, 6.3, 1H, 5'-H_aH_b), 2.56 (t, J = 7.2, 2H, CH₂S), 1.69-1.64 (m, 2H, CH₂CH₂CH₂) 1.40 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, CD₃CN) δ_{C} 158.3 (C),

155.9 (C), 152.8 (CH), 149.1 (C), 139.9 (CH), 115.3 (C), 88.5 (CH), 84.15 (CH), 73.6 (CH), 72.7 (CH), 70.0 (C), 34.0 (CH₂), 34.3 (CH₂), 29.9 (CH₂), 27.7 (CH₃), 25.7 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd C₁₈H₂₉N₆O₅S⁺ 441.1915, found 441.1916.

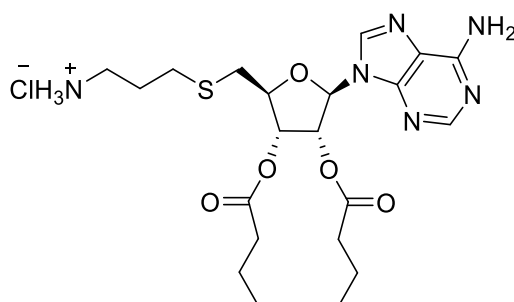
Preparation of (2*S*,3*S*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-(((3-((*tert* butoxycarbonyl)amino)propyl)thio)methyl)tetrahydrofuran-3,4-diyl)dibutyrate **73**



A solution of butyric anhydride (0.41 mL, 2.55 mmol) was added to a solution of **72** (1.80 g, 1.02 mmol) in pyridine (5 mL) at room temperature under argon. The reaction was left for 16 hours at room temperature with stirring. Water (4 mL) was added to the mixture and then extracted with ethyl acetate (4 × 4 mL). The combined organic layers were dried over Na₂SO₄, filtered, concentrated then the residue was purified by silica gel column chromatography using MeOH/CH₂Cl₂ (2/100, v/v) to give the product as white foam (0.33 g, 55 %): $[\alpha]^{22}_D$ -8.31 (c 1.0 in CHCl₃); $\nu_{\max}/\text{cm}^{-1}$ 2973 (C-H), 1747, 1708 and 1632 (C=O), 1246 (C-N), 1168 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_H 8.33 (s, 1H, Ar-H), 8.08 (s, 1H, Ar-H), 6.27 (s, 2H, adenosine-NH₂), 6.16 (d, J = 5.0, 1H, 1'-H), 5.96 (app. t, J = 5.0, 1H, 2'-H), 4.79 (br. s, 1H, NHBoc), 4.34 (dd, J = 5.7, 2.0, 1H, 3'-H) 4.39-4.35 (m, 1H, 4'-H), 3.16 (app. q, 6.2, 2H, CH₂NHBoc), 2.91 (dd, J = 14.2, 5.0, 2H, 5'-H_aH_b-),

2.85 (dd, $J = 14.2, 5.9$, 1H, 5'-H_aH_b), 2.56 (t, $J = 7.3$, 2H, CH₂S), 2.34 (t, $J = 7.6$, 2H, CH₂CO), 2.30 (t, $J = 7.6$, 2H, CH₂CO) 1.69-1.64 (m, 6H, CH₂CH₂CH₂) 1.42 (s, 9H, C(CH₃)₃), 1.0 (t, $J = 7.0$, 3H, CH₃CH₂), 1.00 (t, $J = 7.2$, 3H, CH₃CH₂); ¹³C NMR (100 MHz, CDCl₃) δ_c 172.4 (C), 172.1 (C), 156.0 (C), 155.6 (C), 153.1 (CH), 149.1 (C), 139.9 (CH), 119.3 (C), 86.0 (CH), 82.5 (CH), 77.0 (C), 72.9 (CH), 72.6 (CH), 39.3 (CH₂), 35.9 (CH₂), 35.5 (CH₂), 34.3 (CH₂), 30.5 (CH₂), 29.7 (CH₂), 28.3 (CH₃), 18.3 (CH₂), 18.1 (CH₂), 13.4 (CH₃), 13.3 (CH₃); HRMS (ESI+) m/z : (M+H⁺) calcd C₂₆H₄₁N₆O₇S⁺ 581.2752, found 581.2765.

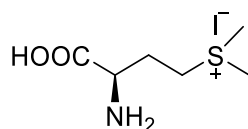
Preparation of 3-((((2*S*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-bis(butyryloxy)tetrahydrofuran-2-yl)methyl)thio)propan-1-aminium chloride **74**



A suspension of **73** (80.0 mg, 0.15 mmol) was dissolved in neat TFA (5 mL) and stirred for 1 hour at room temperature. H₂O (0.5 mL) was then added and stirring continued for 4 hours. The reaction mixture was concentrated and dissolved again in H₂O (5 mL). The mixture was extracted with EtOAc (2 × 4 mL). The organic layers were combined and dried over Na₂SO₄, filtered, concentrated then the residue was purified by silica gel column chromatography using MeOH in NH₃/CH₂Cl₂ (2/100, v/v) to give the product as pale yellow solid (50 mg, 83 %): $[\alpha]^{22}_{\text{D}} -10.2$ (c 1.0 in CHCl₃); $\nu_{\text{max}}/\text{cm}^{-1}$ 2973 (C-H), 1747 and 1708 (C=O), 1246

(C-N), 1168 (C-O); ^1H NMR (400 MHz, CDCl_3) δ_{H} 8.37 (s, 1H, Ar-H), 8.06 (s, 1H, Ar-H), 6.18 (d, $J = 6.0$, 1H, 1'-H), 5.96 (app. t, $J = 6.0$, 1H, 2'-H), 5.85 (s, 2H, adenosine- NH_2), 5.64 (dd, $J = 6.0$, 4.0, 1H, 3'-H) 4.39-4.35 (m, 1H, 4'-H), 3.06 (dd, $J = 14.0$, 5.0, 1H-5' H_aH_b), 2.99 (dd, $J = 14.2$, 5.7, 1H-5' H_aH_b), 2.80 (app. t, $J = 6.72$, 2H, CH_2NH_2), 2.70 (t, 7.31, 2H, CH_2S), 2.38 (t, 7.40, 2H, CH_2CO), 2.74 (t, 7.40, 2H, CH_2CO), 1.91-1.64 (m, 6H, $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.05 (t, 7.5, 3H, CH_3CH_2), 1.0 (t, 7.5, 3H, CH_3CH_2); ^{13}C NMR (100 MHz, CDCl_3) δ_{C} 172.4 (C), 172.1 (C), 155.0 (C), 153.1 (CH), 149.1 (C), 139.9 (CH), 119.3 (C), 86.0 (CH), 82.5 (CH), 72.9 (CH), 72.6 (CH), 40.3 (CH_2), 35.9 (CH_2), 35.5 (CH_2), 34.3 (CH_2), 30.5 (CH_2), 29.7 (CH_2), 28.3 (CH_3) 18.3 (CH_2), 18.1 (CH_2), 13.4 (CH_3), 13.5 (CH_3); HRMS (ESI+) m/z : ($\text{M}+\text{H}^+$) calcd $\text{C}_{21}\text{H}_{32}\text{N}_6\text{O}_5\text{S}^+$ 481.2228, found 481.2232.

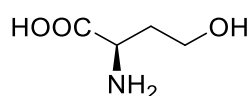
Preparation of methionine methyl sulfonium iodide **76**⁸⁰



L-Methionine (6.00 g, 40.2 mmol) was dissolved in water (80 mL), and a solution of methyl iodide (6.14 mL, 102.0 mmol) was added slowly at room temperature. The mixture was stirred at 40 °C for 16 hours. The solvent was then evaporated under reduced pressure. The resulting solid was dissolved in water (40 mL) and ethanol (120 mL), causing the precipitation of a white crystalline solid. The resulting mixture was allowed to stand for 5 hours to allow further precipitation. The precipitate was collected by filtration to give the product as a white solid (9.86 g, 84 %): m.p. 150-153 °C [lit. 150 °C];⁸⁰ $[\alpha]^{25}_{\text{D}}$ -7.0 (c 1.0 in H_2O) [lit. $[\alpha]^{18}_{\text{D}}$ -8.0 in H_2O]; $\nu_{\text{max}}/\text{cm}^{-1}$ 1609 (C=O); ^1H NMR (400 MHz, D_2O) δ_{H} 3.83 (br. t, $J = 6.2$,

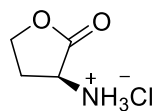
1H, CHCH₂), 3.52-3.48 (m, 1H, CH_aH_bCH), 3.38-3.33 (m, 1H, CH_aH_bCH), 2.91 (s, 6H, 2 x CH₃) 2.27-2.22 (app. br. q, *J*= 7.3, 2H, CHCH₂); ¹³C NMR (100 MHz, D₂O) δ_C 172.4 (C), 52.7 (CH), 39.5 (CH₂), 25.0 (CH₂), 24.7 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₆H₁₅NO₂S⁺ 164.0740, found 164.0758.

Preparation of *S*-homoserine **77**⁸⁰



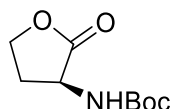
(*S*)-Methionine methyl sulfonium iodide **76** (9.50 g, 32.7 mmol) and a solution of sodium hydrogen carbonate (2.74 g, 32.7 mmol) in water (15 mL) were mixed and heated at reflux for 24 hours. The reaction mixture was then cooled to room temperature. The solvent was removed *in vacuo* to produce a thick syrup. This residue was dissolved in water (7 mL) and heated at 60-70 °C. Acetone (15 mL) was added to the mixture followed by ethanol (150 mL) to give precipitate (*S*)-homoserine as a white solid (3.74, 53 %): m.p.168-171 °C [lit. 199 °C];⁸⁰ [α]²⁴_D + 23 (c 1.0 in H₂O) [lit. [α]¹⁸_D +16 in H₂O];⁸⁰ ν_{max} /cm⁻¹ 3168 and 2949 (br, O-H), 1645 (C=O); ¹H NMR (400 MHz, D₂O) δ_H 3.73 (app. q, *J*= 4.8, 1H, CHCH₂), 3.46-3.40 (m, 2H, CH₂OH) 1.98-1.92 (m, 1H, CHCH_aCH_b), 1.87-1.80 (m, 1H, CHCH_aCH_b); ¹³C NMR (100 MHz, D₂O) δ_C 174.3 (C), 58.6 (CH), 53.3 (CH₂), 32.1 (CH₂); HRMS (ESI+) *m/z*: (M+Na⁺) calcd C₄H₁₀NO₃Na⁺ 142.0480, found 142.0470. The lower melting point of this compound could be affected by impurities when compared with literature value.

Preparation of (S)-homoserine lactone hydrochloride salt **78**⁸⁰



A solution of (S)-homoserine **77** (3.50 g, 29.1 mmol) in 2.4 M aqueous HCl (65 mL) was heated at reflux for 3 hours, then stirred at room temperature for 16 hours. Ethanol (30 mL) was added to the mixture, and most of the solvent was removed *in vacuo*. Then, the residue was cooled in the fridge for one hour. The precipitate, which appeared during this period was collected by filtration and washed with cold ethanol to give (S)-homoserine lactone hydrochloride salt as colourless crystals (1.70 g, 57 %): m.p. 235-238 °C [lit. 200 °C];⁸⁰ $[\alpha]^{23}_{\text{D}}$ -10.94 (c 1.0 in H₂O) [lit. $[\alpha]^{18}_{\text{D}}$ -26.0 in H₂O];⁸⁰ $\nu_{\text{max}}/\text{cm}^{-1}$ 1785 (C=O); ¹H NMR (400 MHz, D₂O) δ_{H} 4.58 (t, d J = 9.2, 1.2, 1H, $\text{CH}_a\text{H}_b\text{O}$), 4.39-4.32 (m, 2H, $\text{CH}_a\text{H}_b\text{O}$, NHCH), 2.80 (dddd, J = 12.7, 8.9, 6.0, 1.2 1H, $\text{CH}_a\text{H}_b\text{CH}$), 2.36 (dddd, J = 12.7, 11.7, 11.1, 9.2, 1H, $\text{CH}_a\text{H}_b\text{CH}$); ¹³C NMR (100 MHz, D₂O) δ_{C} 174.5 (C), 67.4 (CH₂), 48.5 (CH), 26.78 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd C₄H₈NO₂⁺ 102.0550, found 102.0543.

Preparation of N-Boc-S-homoserine lactone **79**

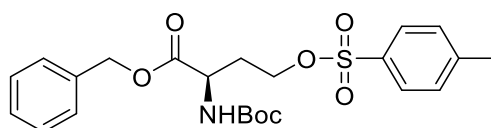


A solution of S-homoserine lactone HCl (200 mg, 1.44 mmol) in DMF (2 mL) was treated with Et₃N (0.11 mL) and di-*tert*-butyl dicarbonate (380 mg, 1.73 mmol) at 0 °C. The reaction mixture was stirred for 2 hours at 0 °C then warmed to room temperature. The solvent was removed *in vacuo*, and the residue was dissolved

in CH₂Cl₂ (12 mL). The mixture was washed with saturated aqueous KHSO₄ (3 × 12 mL) and brine (1 × 12 mL). The combined organic layers were dried over sodium sulfate, and the organic layer was concentrated. The residue was purified by flash chromatography using CH₂Cl₂/MeOH (98 / 2 : v/v) to give product as a white foam (270 mg, 69 %): m.p. 97-99 °C; $[\alpha]^{23}_{\text{D}} +5.12$ (c 1.0 in CHCl₃); $\nu_{\text{max}}/\text{cm}^{-1}$ 1784 and 1714 (C=O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 5.15 (br. s, 1H, NH) 4.44 (t, d $J = 9.2, 1.2$, 1H, CH_aH_bO), 4.26-4.20 (m, 2H, CH_aH_bO, NHCH), 2.77 (m, 1H, CH_aH_bCH), 2.25 (m, 1H, CH_aH_bCH), 1.46 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, D₂O) δ_{C} 175.3 (C), 80.1 (C), 65.7 (CH₂), 50.1 (CH), 30.45 (CH₂), 28.2 (CH₃); HRMS (ESI+) m/z : (M+Na⁺) calcd C₉H₁₅NO₄Na⁺ 224.0899, found 224.0893.

Preparation of benzyl *N*-(*tert*-butoxycarbonyl)-*O*-tosyl-*S*-homoserinate

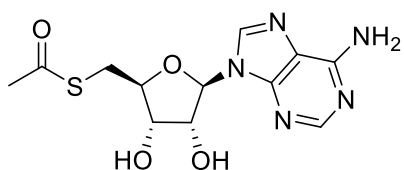
82



Pyridine (1.34 mL) was added over 5 minutes to a solution of *N*-Boc-*S*-homoserine benzyl ester **35** (1.25 g, 4.04 mmol) in the presence of *p*-toluenesulfonyl chloride (1.15 g, 6.06 mmol) in dry CH₂Cl₂ (8 mL) at 0 °C. The mixture was stirred for 16 hours at room temperature under argon. The mixture reaction was quenched by addition of saturated aqueous citric acid and extracted with CH₂Cl₂ (3 × 15 mL). The organic layers were combined and dried over Na₂SO₄. The solution was evaporated *in vacuo* to give a yellow oil, which was purified by chromatography with MeOH/CH₂Cl₂ (1/99, v/v) as eluent to provide a white solid (1.37 g, 80 %): m.p. 102-104 °C; $[\alpha]^{23}_{\text{D}} -28$ (c 1.0 in CHCl₃); ¹H NMR (400 MHz, CDCl₃) δ_{H} 7.96 (d, $J = 8.3$, 2H, Ar-H), 7.77 (d, $J = 8.3$, 2H, Ar-H), 7.38-7.32 (m, 5H, Ph), 5.23 (d, $J =$

6.6, 1H, NHBoc), 5.15 (s, 2H, CH₂Ph), 4.35-4.30 (m, 1H, CHNHBoc), 4.11 (t, *J*= 5.9, 2H, CH₂OTs), 2.45 (s, 3H, CH₃), 2.39-2.33 (m, 1H, CH_aH_bCH), 2.34-2.30 (m, 1H, CH_aH_bCH), 1.45 (s 9H, C(CH₃)₃); ¹³C NMR (100MHz, CDCl₃) δ_c 171.4 (C), 155.1 (C), 144.2 (C), 135.1 (C), 132.4 (C), 130.4 (CH), 130.0 (CH), 129.1 (CH), 128.9 (CH), 127.9 (CH), 127.0 (CH), 80.1 (C), 67.5 (CH₂), 66.2 (CH₂), 50.6 (CH), 31.50 (CH₂), 28.2 (CH₃), 21.8 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₃H₂₉NO₇Na⁺ 486.1562, found 486.1576.

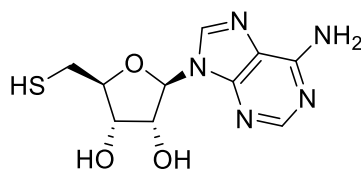
Preparation of *S*-(((2*S*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)ethanethioate **83**⁷⁸



A solution of 5-acetylthio-5-deoxy-2',3'-O-isopropylideneadenosine **65** (700 mg, 1.91 mmol) in neat TFA (25 mL) was stirred for 1 hour at room temperature. Then water (2.5 mL) was added and stirred for 2 hours. The solvent was removed under reduce pressure. The residue was quenched by addition of Na₂CO₃ (10 mL), extracted with ethyl acetate (3 × 20 mL). The combined organic layer were dried over magnesium sulfate, filtered and concentrated to produce product as a white solid (560, mg, 90 %): m.p.105-107 °C; [α]²²_D +10.7 (c 1.0 in CHCl₃); ν_{max}/cm⁻¹ 2094 (C-H), 1706 (C=O), 1240 (C-N), 1100 (C-O); ¹H NMR (400 MHz, CD₃OD) δ_H 8.36 (s, 1H, adenosine-*H*), 8.17 (s, 1H, adenosine-*H*), 5.89 (d, *J*= 5.9, 1H, 1'-*H*), 5.54 (br. s, 1H, OH), 5.40 (br. s, 1H, OH), 4.79 (app. t, *J*= 5.0, 1H, 2'-*H*), 4.12 (app. t, *J*= 5.0, 1H, 3'-*H*), 4.15-4.10 (m, 1H, 4'-*H*), 3.38 (dd, *J*= 13.8, 5.7, 1H, 5'-*H*_aH_b), 3.20 (dd, *J*= 13.8, 7.5, 1H, 5'-*H*_aH_b), 2.34 (s, 3H, CH₃CO); ¹³C NMR (100

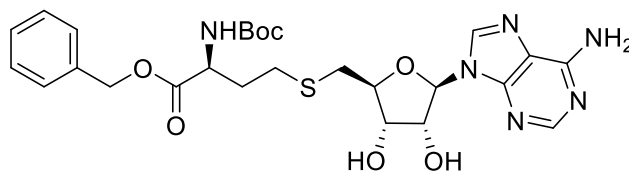
MHz, CD₃OD) δ_c 195.1 (C), 155.9 (CH), 152.4 (C), 149.2 (C), 140.0 (CH), 119.0 (C), 88.7 (CH), 83.33 (CH), 73.3 (CH), 72.77 (CH), 30.8 (CH₂), 28.99 (CH₃); HRMS (ESI+) m/z : (M+H⁺) calcd C₁₂H₁₆N₅O₄S⁺ 326.0918, found 326.0914. The literature does not show the $[\alpha]_D$ and melting point.⁷⁸

Preparation of (2*S*,3*S*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-(mercapto-methyl)tetrahydrofuran-3,4-diol **84**⁷⁸



A suspension of 5' ethanethioate-5'-deoxyadenosine **83** (50.0 mg, 0.15 mmol) in methanol saturated with ammonia (3 mL) was stirred for 30 minutes at room temperature under argon. The solvent was evaporated under reduce pressure to give product as pale-yellow oil (45 mg, 90 %): The suspension was used directly in the next step without any further purification, $[\alpha]^{22}_D$ +10.7 (c 1.0 in CHCl₃); $\nu_{\max}/\text{cm}^{-1}$ 3414 (OH), 2991 (C-H), 1706, 1240 (C-N); ¹H NMR (400 MHz, CH₃Cl₃) δ_H 8.28 (s, 1H, adenosine-*H*), 8.22 (s, 1H, adenosine-*H*), 6.00 (d, J = 5.4, 1H, 1'-*H*), 4.85 (app. t, J = 5.3, 1H, 2'-*H*), 4.25 (t, J = 5.3, 1H, 3'-*H*), 4.15-4.11 (m, 1H, 4'-*H*), 3.39 (dd, J = 14.0, 6.0, 1H, 5'-*H_aH_b*), 3.28 (dd, J = 14.0, 6.9, 1H, 5'-*H_aH_b*); ¹³C NMR (100 MHz, CH₃Cl₃) δ_c 156.1 (C), 153.0 (CH), 149.0 (C), 140.0 (CH), 120.7 (C), 90.5 (CH), 86.1 (CH), 84.0 (CH), 83.0 (CH), 41.0 (CH₂); HRMS (ESI+) m/z : (M+H⁺) calcd C₁₀H₁₄N₅O₃S⁺ 284.0812 found 284.0801. The literature does not show the $[\alpha]_D$ and melting point.⁷⁸

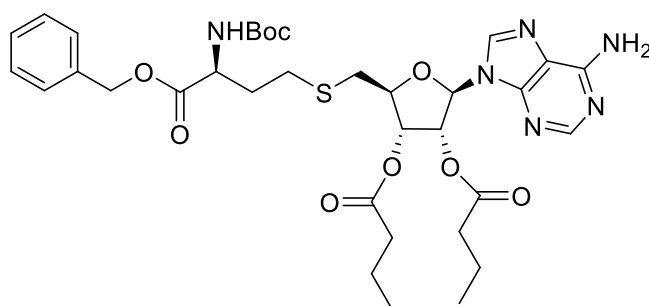
Preparation of *tert*-butyl ((*S*)-4-(((2*S*,3*S*,4*R*,5*R*)-5-(6-amino-9*H*-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl) thio)-1-(benzyloxy)-1-oxobutan-2-yl)- λ^2 -azanecarboxylate **85**



A suspension of *N-tert*-Butyloxycarbonyl-O-tosyl-*S*-homoserine benzyl ester **82** (0.32 g, 0.70 mmol) in DMF (2 mL) was added to a compound **84** (0.20 g, 0.70 mmol) and potassium carbonate (0.19 g, 1.4 mmol) in DMF (5 mL) at room temperature under argon. Sodium borohydride (0.05 g, 1.4 mmol) was then added to reduce the disulfide bond of starting material. The reaction mixture was heated to 100 °C for 30 minutes and left stirring at room temperature for one hour. Water (4 mL) was added, and the mixture was extracted with ethyl acetate (3 × 5 mL). The organic layers were washed with water (5 mL) and brine (5 mL), dried over Na₂SO₄, filtered and evaporated to dryness, The residue was purified by silica gel chromatography using methanol/dichloromethane (2/100, v/v) as eluent to give the product as pale yellow oil (0.25 g, 71 %): $[\alpha]^{23}_D$ -24.0 (c 1.0 in CHCl₃); $\nu_{\max}/\text{cm}^{-1}$ 3414 (N-H), 3069 (C-H, aromatic), 2984 (C-H), 1731 and 1710 (C=O), 1248 (C-N), 1091 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_H 8.23 (s, 1H, Ar-*H*), 8.08 (s, 1H, Ar-*H*), 7.37 (m, 5H, *Ph*), 6.23 (s, 2H, adenosine-NH₂), 5.95 (d, *J*= 4.8, 1H, 1'-*H*), 5.74 (m, 1H, NHBoc) 5.20 (s, 2H, CH₂Ph), 4.76 (app. t, *J*= 4.8, 1H, 2'-*H*), 4.33 (app. t, *J*= 4.8, 1H, 3'-*H*), 4.28-4.24 (m, 1H, 4'-*H*), 4.14-4.10 (m, 1H, CHNHBoc), 2.98-2.95 (m, 2H, CH₂S), 2.68-2.64 (dd, *J*= 13.0, 6.0, 1H, 5'-*H_aH_b*), 2.54-2.50 (dd, *J*= 13.0, 6.0, 1H, 5'-*H_aH_b*), 1.98-1.95 (m, 1H, CH_aCH_bCH), 1.89-1.85 (m, 1H, CH_aCH_bCH) 1.41 (s, 9H, C(CH₃)₃); ¹³C NMR (100 MHz, CDCl₃) δ_C 176 (C), 155.2 (C), 155.9 (C), 153.1 (CH), 150.1 (C), 139.9 (CH), 135 (C), 126.3

(CH), 126.0 (CH), 125.0 (CH), 119.5 (C), 87.7 (CH), 84.7 (CH), 79.6 (C), 74.7 (CH), 73.6 (CH), 67.4 (CH₂), 54.4 (CH₂), 53.7 (CH), 31.3 (CH₂), 29.3 (CH₂), 27.0 (CH₃); HRMS (ESI+) m/z : (M+H⁺) calcd C₂₆H₃₄N₆O₇S⁺ require 574.2020, found 574.2029.

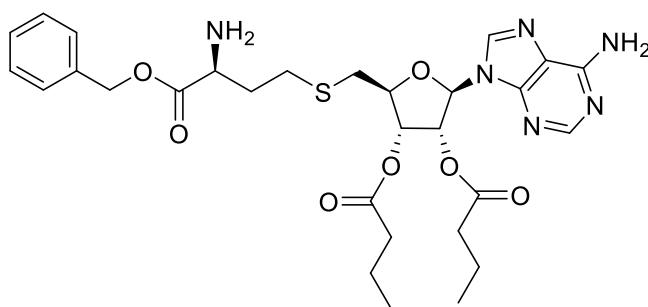
Preparation of (2*S*,3*S*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((((*S*)-4-(benzyloxy)-3-((*tert*-butoxycarbonyl)amino)-4-oxobutyl)thio)methyl)tetrahydrofuran-3,4-diyl dibutyrate **86**



A solution of butyric anhydride (0.21 mL, 1.36 mmol) was added to a solution of **85** (0.20 g, 0.34 mmol) in pyridine (5 mL) at room temperature under argon. The reaction was left for 16 hours at room temperature with stirring. Water (4 mL) was added and the mixture extracted with ethyl acetate (4 × 4 mL). The combined organic layers were dried over Na₂SO₄, filtered, concentrated then the residue was purified by silica gel column chromatography using MeOH/CH₂Cl₂ (2/100, v/v) to give the product as white foam (0.21 g, 87 %): $[\alpha]^{23}_{\text{D}}$ -8.0 (c 1.0 in CHCl₃); $\nu_{\text{max}}/\text{cm}^{-1}$ 3414 (N-H), 2973 (C-H), 1708 and 17.00 (C=O), 1248 (C-N), 1084 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.28 (s, 1H, Ar-*H*), 8.08 (s, 1H, Ar-*H*), 7.38 (m, 5H, *Ph*), 6.13 (d, J = 5.4, 1H, 1'-*H*), 6.11 (s, 2H, adenosine-NH₂), 6.06 (app. t, J = 5.4, 1H, 2'-*H*), 5.65 (m, 2H, NHBoc and 3'-*H*), 5.14 (dd, J = 12.0, 5.0, 2H, CH₂Ph),

4.34-4.30 (m, 1H, 4'-H), 4.26-4.22 (m, 1H, CHNH₂Boc), 3.00-2.95 (m, 2H, CH₂S), 2.61 (dd, *J* = 13.0, 6.0, 1H, 5'-H_aH_b), 2.54 (dd, *J* = 13.0, 6.0, 1H, 5'-H_aH_b), 2.39 (t, *J* = 7.3, 2H, CH₂O), 2.29 (t, *J* = 7.3, 2H, CH₂O) 1.66-1.60 (m, 3H, CH_aCH_bCH and CH₂CH₃), 1.57-1.53 (m, 3H, CH_aCH_bCH and CH₂CH₃) 1.41 (s, 9H, C(CH₃)₃), 0.98 (t, *J* = 7.3, CH₃CH₂), 0.90 (t, *J* = 7.3, CH₃CH₂); ¹³C NMR (100 MHz, CDCl₃) δ_C 172.9 (C), 172.5 (C), 172.1 (C), 155.9 (C), 155.2 (C), 153.1 (CH), 150.1 (C), 139.9 (CH), 136.1 (C), 128.3 (CH), 128.0 (CH), 127.0 (CH), 119.5 (C), 87.7 (CH), 84.7 (CH), 79.6 (C), 74.7 (CH), 73.6 (CH), 67.4 (CH₂), 53.7 (CH), 53.4 (CH₂), 31.3 (CH₂), 31.1 (CH₂), 29.3 (CH₂), 29.3 (CH₂), 27.0 (CH₃) 18.3 (CH₂), 18.0 (CH₂), 13.9 (CH₃), 13.0 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₃₄H₄₆N₆O₉S⁺ require 715.31020, found 715.3108.

Preparation of (2*S*,3*S*,4*R*,5*R*)-2-((((*S*)-3-amino-4-(benzyloxy)-4-oxobutyl)thio)methyl)-5-(6-amino-9*H*-purin-9-yl)tetrahydrofuran-3,4-diyl dibutyrate **87**



A suspension of **86** (100.0 mg, 0.13 mmol) was dissolved in neat TFA (3 mL) and stirred for 1 hour at room temperature. H₂O (0.5 mL) was then added and stirring continued for 4 hours. The reaction mixture was concentrated and dissolved again in H₂O (3 mL) and washed with EtOAc (2 × 4 mL). The combined organic layers

were dried over Na₂SO₄, filtered, concentrated then the residue was purified by silica gel column chromatography using MeOH in NH₃/CH₂Cl₂ (2/100, v/v) to give the product as pale yellow solid (60 mg, 83 %): [α]²³_D -6.8 (c 1.0 in CHCl₃); ν_{max} /cm⁻¹ 3413 (N-H), 2969 (C-H), 1742 and 1632 (C=O), 1241 (C-N), 1091 (C-O); ¹H NMR (400 MHz, CDCl₃) δ_{H} 8.26 (s, 1H, Ar-*H*), 8.08 (s, 1H, Ar-*H*), 7.38 (m, 5H, *Ph*) 6.13 (d, *J*= 5.4, 1H, 1'-*H*), 6.11 (s, 2H, adenosine-NH₂), 5.66 (app. t, *J*= 5.4, 1H, 2'-*H*), 5.14 (m, 3H, CH₂Ph and 3'-*H*), 4.34-4.30 (m, 1H, 4'-*H*), 3.55-3.51 (m, 1H, CHNH₂), 3.09 (dd, *J*= 14.0, 7.0, 2H, 5'-*H*_a*H*_b and 5'-*H*_a*H*_b), 3.00-2.95 (m, 2H, CH₂S), 2.39 (t, *J*= 7.3, 2H, CH₂O), 2.28 (t, *J*= 7.3, 2H, CH₂O), 1.67-1.63 (m, 3H, CH_aCH_bCH and CH₂CH₃), 1.57-1.53 (m, 3H, CH_aCH_bCH and CH₂CH₃), 0.98 (t, *J*= 7.3, CH₃CH₂), 0.80 (t, *J*= 7.3, CH₃CH₂); ¹³C NMR (100 MHz, CDCl₃) δ_{C} 172.9 (C), 172.5 (C), 172.1 (C), 155.9 (C), 153.1 (CH), 150.1 (C), 139.9 (CH), 136.1 (C), 128.3 (CH), 128.0 (CH), 127.0 (CH), 119.5 (C), 87.7 (CH), 84.7 (CH), 74.7 (CH), 73.6 (CH), 67.4 (CH₂), 53.4 (CH), 32.6 (CH₂), 31.3 (CH₂), 31.1 (CH₂), 29.3 (CH₂), 28.3 (CH₂), 27.0 (CH₃) 18.3 (CH₂), 18.0 (CH₂), 13.9 (CH₃), 13.0 (CH₃); HRMS (ESI+) *m/z*: (M+H⁺) calcd C₂₉H₃₈N₆O₇S⁺ 615.2595, found 615.2604.

7.2. Materials and methods of biological chapters

7.2.1. Chemical components

All chemicals used in this thesis were provided by Sigma-Aldrich (UK) unless stated otherwise.

7.2.2. Buffers pH adjustment

All buffers used and their compositions are listed in Table 7.1. All water used was deionised using a 0.22 µm membrane filter through *Milli-Q* water purifiers and autoclaved prior to use.

The pH calibration involves cleaning the Bluelab® Leap™ pH Probe tip and then calibrating in two solution, either pH 4.00 or pH 10.00. Firstly, rinse the pH electrode with distilled water. Secondly, calibrate the pH electrode with pH 7.0 then pH 4.0 or pH 10.0. If a reading below pH 7.0, the pH electrode can calibrate using pH 7.0 and pH 4.0 calibration solutions. If a reading above pH 7.0, the pH electrode can calibrate using pH 7.0 and pH 10.0 calibration solutions.

Buffer pH was measured using a pH meter (desktop-pH-meter five easy (TM) F20-meter incl) and adjusted in 800 mL batches then made up to 1 litre by addition of deionised water. pH was adjusted by careful addition of either 2.0 M aqueous sodium hydroxide solution or 2.0 M hydrochloric acid to the stirring buffer until a stable desired pH was obtained.

Solution	Chemical components
TAE Buffer	50 mM EDTA, 50 mM acetic acid, 40 mM Tris-HCl, pH 8.0.
Cracking buffer	50 % (w/v) sucrose, 10 % (w/v) SDS, 50 mM NaOH, and 25-50 ng bromophenol blue in H ₂ O
Cell lysis buffer	300 mM NaCl, 10 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0
chaperonin removal buffer	150 mM NaCl, 10 mM MgCl ₂ , 5 mM ATP and 20 mM HEPES-NaOH (pH 7.0)
His Trap binding buffer	300 mM NaCl, 10 % (v/v) glycerol, 10 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0
His Trap elution buffer	300 mM NaCl, 10 % (v/v) glycerol, 500 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0
Gel filtration buffer	150-300 mM NaCl, 2-5 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0
SDS running buffer	4 % (w/v) Glycine, 0.1 % (w/v) SDS, 125 mM Tris-HCl and pH 6.8
SDS loading dye buffer	160 mM Tris-HCl (pH 6.8), 5 % (w/v) SDS, 30 % (v/v) glycerol, 20 mM DTT, 0.03 % (w/v) bromophenol blue
Resolving gel of SDS PAGE composition	3.7 mL of H ₂ O, 8.0 mL of 30 % Acrylamide/Bis-Acrylamide (v/v), 4.0 mL of 500 mM Tris-HCl (pH 6.6), 160 µL of 10 % (w/v) SDS, 160 µL of 10 % (w/v) APS, 16.0 µL of TEMED
Stacking gel of SDS PAGE composition	5.4 mL of H ₂ O, 2 mL of 30% Acrylamide/Bis-Acrylamide (v/v), 2.5 mL of 500 mM Tris-HCl (pH 6.8), 100 µL of 10 % (w/v) SDS, 100 µL of 10 % (w/v) APS, 10 µL of TEMED
Transfer buffer	192 mM glycine, 20 % (v/v) MeOH, 25 mM Tris-HCl (pH 8.3)
Blocking buffer	150 mM NaCl, 0.1 % (v/v) TWEEN-20, 5 % (w/v) skimmed milk powder and 50 mM Tris-HCl (pH 7.5),
ECL buffer	22 µL coumarin, 50 µL luminol and 3 µL hydrogen peroxide made up to 10 mL with 0.1 M Tris-HCl (pH 8.5)
ITC buffer	150 mM NaCl, 1 % (v/v) glycerol, 20 mM Tris-HCl and pH 8.0
LB agar	agar 15 g/L, yeast extract 5 g/L, tryptone 10 g/L and sodium chloride 10 g/L
LB medium	10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride

Table 7-1: List of buffers used during biological work with individual component concentrations specified. All buffers were made using deionised water.

7.2.3. Biological material

All molecular biology consumables were supplied by Thermo Fisher Scientific unless otherwise specified.

DNA constructs coding for His-tagged PRMT1 in pET21a or CARM1 in pMAL-c2xE (Table 7-2) were supplied by Dr Emma Gunnell (University of Nottingham). Expression constructs coding for hPRMT2 and hPRMT6 (Table 7-2) were kindly provided by Dr Clare Davies (College of Medical and Dental Sciences, University of Birmingham). Vectors used to clone sequences of target proteins (Table 7-2) were supplied by Dr Dreveny's group (University of Nottingham).

7.3. Molecular biology techniques

7.3.1. Sterilisation and sterile technique

To reduce the risk of contamination, all glassware used for buffer and cell preparation was autoclaved by pressurised saturated steam, 121 °C at 100 kPa for 20 mins. Stock solutions and buffers were sterilised by filtration using a 0.2 µm syringe filter (Thermo Fisher Scientific). All centrifuge tubes and Petri dishes were sterilised from commercial suppliers unless otherwise stated. All water used was deionised and autoclaved prior to use. All work involving live cells was carried out using proper aseptic technique.

7.3.2. Agar plates

Bacterial suspensions were spread on Lennox Luria-Bertani media agar plates, which contained the appropriate antibiotic in conjunction with the selection method of the vector in question in order to obtain individual bacterial colonies. 40 g of LB agar (Lennox) was dissolved in 1-litre of deionised water and autoclaved. LB agar contained 15 g/L agar, 5 g/L yeast extract, 10 g/L tryptone and 10 g/L sodium chloride, respectively dissolved in deionised water. When required, solid agar was melted in a microwave and left to cool to approximately 50 °C. 10 µL of appropriate antibiotic (kanamycin (0.050 mg/mL), chloramphenicol (0.035 mg/mL), or ampicillin (0.1 mg/mL)) was added into 10 mL of LB agar, by using conical centrifuge tubes (500/CS, 50 mL) (Sigma-Aldrich, UK) and then poured into Petri dishes and allowed to set at room temperature, next to a blue flame. Plates were either used immediately in assays looking for selective growth post bacterial transformation or stored at 4 °C until used.

7.3.3. Plasmid DNA purification

Plasmids were purified using either a GenElute PCR Clean-Up kit or a GenElute Gel Extraction Kit or a GenElute Plasmid Miniprep Kit (all Sigma-Aldrich, UK). The basis for these kits is the alkaline lysis of *E. coli* cells, followed by the binding of DNA to a silica membrane with samples processed as described in the kit protocol. The DNA plasmid concentrations were determined by the use of a Nanodrop as described further in section 7.12.1.

7.3.4. DNA agarose gel electrophoresis

For DNA separation, samples were run on a 1 % (w/v) agarose gel for 40 min at 100 V. Preparing the agarose gel was as follows: 0.5 g of agarose (Thermo Fisher Scientific) was dissolved in 50 mL of TAE buffer containing (50 mM EDTA, 50 mM acetic acid, 40 mM Tris-HCl and pH 8.0). The mixture was heated in a microwave until fully dissolved and briefly allowed to cool. The mixture was then poured into a gel caster (BioRad Sub-cell GT UV-Transparent Mini-Gel Tray) fitted with the appropriate well comb and stained using Nancy-520 (Sigma Aldrich, UK) to visualise bands.

Samples were prepared as follows: 2 μ L plasmid DNA was mixed with 2 μ L of 6 $\times$ DNA Loading Dye (Thermo Fisher Scientific), and 6 μ L of water were then loaded on the DNA-agarose gel stain. Reaction products were confirmed estimating the size of bands by comparing them with bands present in standards such as the 100 bp and 1kb DNA ladders (NEB). Gels were imaged using a Dark Reader Transilluminator (Clare Chemical).

7.3.5. Glycerol stocks

500 μ L of *E. coli* cultures were grown overnight (see Section 7.6.) and was then resuspended with an equal volume of sterile 30 % (v/v) glycerol, in a sterile polypropylene tube (Thermo Fisher Scientific). Cells suspensions were rapidly frozen using liquid nitrogen and stored at - 80 °C, until subsequent use.

7.3.6. Generation of construct

All biological experiments are presented as general procedures below. Where appropriate, specific variations relating to each target protein are listed in tables in the same order that they appear in the thesis.

PRMT2, CARM1, PRMT6 and substrate constructs were utilised, cloned, expressed and purified as summarised in Table 7-2 and illustrated in Figure 7-1. All experiments were carried out by myself, except where indicated.

Chapter	Name of construct	Residue range	Expression Vector	Restriction sites and buffer used	Purification Tag	Cloned by other
3	H4	1-21	pET21a	<i>Bam</i> HI <i>Xho</i> I CutSmart Buffer	N-term GB C-term His	Emma Gunnell
	PRMT1	22-361	pET26b	<i>Xho</i> I <i>Nde</i> I CutSmart Buffer	N-term His	Emma Gunnell
	MBP-CARM1	135-479	pMAL-c2x E	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term MBP C-term His	Emma Gunnell
	2GKG-CARM1	135-479	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 2GKG C-term His	
	2GKG-CARM1	135-479	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 2GKG with TEV cleavage site C-term His	
4	PRMT6	28-375	pPROEX HTb	<i>Bam</i> HI <i>Hind</i> III CutSmart Buffer	C-term His	
	MBP-PRMT6	28-375	pMAL-c2x E	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term MBP C-term His	
	2GKG-PRMT6	28-375	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 2GKG C-term His	
	1BKR-PRMT6	28-375	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 1BKR C-term His	
	PRMT6	39-375	pPROEX HTb	<i>Bam</i> HI <i>Hind</i> III CutSmart Buffer	C-term His	
	MBP-PRMT6	39-375	pMAL-c2x E	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term MBP C-term His	
	2GKG-PRMT6	39-375	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 2GKG C-term His	
	1BKR-PRMT6	39-375	pRSF-13	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term 1BKR C-term His	
	H3	1-21	pET21a	<i>Bam</i> HI <i>Xho</i> I CutSmart Buffer	N-term GB C-term His	
5	RSF1	110-124	pET21a	<i>Bam</i> HI <i>Xho</i> I CutSmart Buffer	N-term GB C-term His	
	FL-PRMT2	1-433	pColdI	<i>Nde</i> I <i>Hind</i> III CutSmart Buffer	N-term His	
	FL-PRMT2	1-433	pColdI	<i>Nde</i> I <i>Hind</i> III CutSmart Buffer	N-term TF	
	MBP-PRMT2	95-433	pMAL-c2x E	<i>Not</i> I <i>Hind</i> III CutSmart Buffer	N-term MBP C-term His	

	2GKG-PRMT2	95-433	pRSF-13	NotI HindIII CutSmart Buffer	N-term 2GKG C-term His	
	1BKR-PRMT2	95-433	pRSF-13	NotI HindIII CutSmart Buffer	N-term 1BKR C-term His	

Table 7-2: Summary of constructs generated

This table shows how each construct was produced, including the residue range for each protein, the expression vectors, restriction enzymes and purification tag use. All experiments were conducted by myself unless carried out by Dr Emma Gunnell, as indicated.

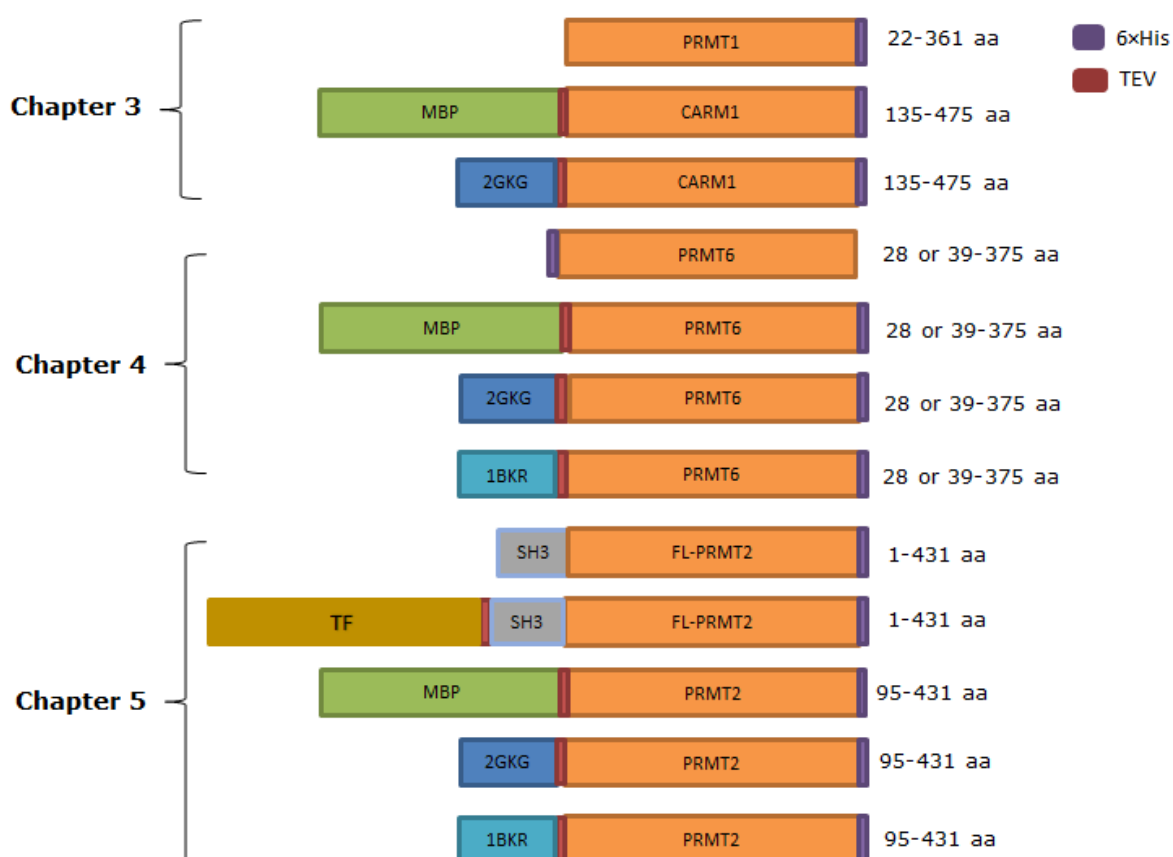


Figure 7-1: Schematic representation of the different catalytic domain constructs of PRMT1, PRMT2, CARM1 and PRMT6 that were cloned. Different colours are indicated for catalytic domains and tags.

7.4. Polymerase chain reaction

PRMT2, CARM1 and PRMT6 constructs were cloned using standard PCR from full-length sequence templates. The primers used to include restriction enzyme sites are included in Table 7-2. The forward and reverse PCR primers used during the cloning of each construct are given in Table 7-3.

Chapter	Name of Construct	Residue range	Forward primer 5' to 3'	Reverse primer 5' to 3'
3	2GKG-CARM1	135-479	GGTACACACCACCAC- CACCACCACTGA CTCGAG CGG	CCGCTCGAGTCAG- TGGTGGTGGTGGTGGTGTG TACC
4	PRMT6	28-375	CGCGGATCCGAGCGG- GAGGCGGCC	ACCAAAGACTTT- GCCATGGAGGACTGAAA- GCTTGGG
	MBP-PRMT6 2GKG-PRMT6 1BKR-PRMT6	28-375	GCGGCCGCGAGAGCGG- GAG GCGGCC	ACCAAAGACTTT- GCCATGGAGGACTGAAA- GCTTGGG
	MBP-PRMT6 2GKG-PRMT6 1BKR-PRMT6	39-375	GCGGCCGCAACTAA- GCGGGAACGGGACCAG	CCCAAGCTTTCAG- TCCTCCATGG- CAAAGTCTTTGGT
4 and 5	H3	1-22	CGCGGATCCGAAAAC- CTGTATTTTCAGGCCCG- TACCAAGCAGACCGC	CCGCTCGAGTTAG- TGGTGGTG GTGGTGGTGGTGGGCCAG CTGCTTGCAGGGG
5	RSF1	110-124	GATCCAAC- CGCTCGTTTAAC TCG- TATGGCGGCGGCCGCC GTATAACCGC	TCGAGGCGGTTATAC- GGGCGGCCGCCGCCATAC GAGTAAA CGAGCGGTTG
	FL-PRMT2	1-433	CTG- GAAGTGCTGTTTCAGGGC CCGATGG- CAACATCAGGTGACTGTC	CCCAA- GCTTTCATCTCCAGATGGG- GAAGAC
	MBP-PRMT2 2GKG-PRMT2 1BKR-PRMT2	95-433	GCGGCCGCGACCCGAG- GACACGTGGCAGG	CCCAA- GCTTTCATCTCCAGATGGG- GAAGACTTTTCTCC

Table 7-3: Different primers used for producing PCR products of different constructs.

Polymerase chain reaction (PCR) was used to amplify the desired section of the gene. Reaction mixtures were made for each PCR reaction (Table 7-4).

Reagent	50 μ L volume	Final amount or concentration
Template plasmid	0.5 μ L	50 – 150 ng of DNA
5 $\times$ Phusion HF buffer	10 μ L	$\times$ 1
dNTPs	1.0 μ L	200 μ M
Phusion DNA polymerase	0.5 μ L	1.0 units/50 μ l PCR
DMSO	2 μ L	2 % (v/v)
Forward primer	2 μ L	20 μ M
Reverse primer	2 μ L	20 μ M
H ₂ O	32 μ L	

Table 7-4: Reagents used in 50 μ L PCR reactions.

The PCR reaction was then carried out using conditions listed in Table 7-5, using a PTC-200 Peltier Thermocycler. All reagent used for the PCR reaction were purchased from Thermo Fisher Scientific.

PCR process	Temperature	Time	cycles
Initial denaturation	95 $^{\circ}$ C	5 minutes	25
Denaturation	95 $^{\circ}$ C	30 minutes	
Annealing	55/65 $^{\circ}$ C	30 seconds	
Extension	72 $^{\circ}$ C	1 minute/1kb	
Final extension	72 $^{\circ}$ C	10 minutes	

Table 7-5: Thermocycler settings for the PCR reaction.

Two annealing temperatures were performed; one reaction at 55 $^{\circ}$ C and other at 65 $^{\circ}$ C to maximise the chance of a successful PCR reaction.

Following PCR, the PCR product was run on an agarose gel electrophoresis (using the protocol described in Section 7.3.4) to confirm the molecular weight and

assess the purity of the PCR product. The size of the PCR product was compared against a standard 100 bp or 1kbp DNA ladder (New England Biolabs (NEB)). Prior to ligation, the pure PCR product was purified using a GenElute PCR clean-up kit (Sigma-Aldrich, UK) in order to remove the PCR primers and Phusion polymerase. In cases where impure PCR products were present, the DNA band corresponding to the PCR product was cut-out and purified using a GenElute Gel Extraction Kit (Sigma-Aldrich, UK), following the instructions provided by Sigma Aldrich.

7.4.1. Restriction enzyme digestion

PCR products and vectors were digested with the appropriate restriction enzymes (Table 7-2) such as *NotI* and *HindIII* (New England Biolabs (NEB)) to produce complementary “sticky ends”, these sticky ends would then improve ligation efficiency in subsequent experiments. Initially, the digestion used 0.5 μL (10 U) of each appropriate restriction enzyme and 2 μL (1 X) of 10 $\times$ CutSmart buffer (New England Biolabs (NEB)) to 17 μL (10 ng/ μL) of DNA (purified PCR product) and was incubated for 16 at room temperature. The digested PCR products were purified using a GenElute PCR clean-up kit (Sigma-Aldrich, UK) to ensure removal of restriction enzymes. The 50 ng/ μL vector was subjected to restriction digestion in the same way as the PCR product.

7.4.2. Phosphatase treatment

The digested plasmid was dephosphorylated at the 5' end using Antarctic Phosphatase (New England Biolabs (NEB)) to prevent self-annealing. The digested reaction mixture 17 μL (2.5 ng/ μL) was dephosphorylated by the addition of 1 μL

(5 U) of Antarctic phosphatase (New England Biolabs (NEB)) and 2 μL of 10 $\times$ Antarctic phosphatase buffer (1 X final concentration) (New England Biolabs (NEB)) and incubated for 2 h at room temperature.

7.4.3. Ligation

DNA ligase was used to form phosphodiester bonds between the 3'-hydroxyl end of one nucleotide with the 5'-phosphate end of another. The result of this process was to join vector with DNA fragment covalently. Each ligation reaction was performed by mixing a DNA fragment with a vector (both pre-digested with the corresponding restriction enzymes) at a ratio of between 3:1 and 6:1, respectively. The ligation reaction contained 1 μL (20 U) of T4 DNA Ligase (New England Biolabs (NEB)), 1 μL of 1 X DNA ligase buffer (NEB), 1 μL (2.5 ng/ μL) of vector and 3 μL or 6 μL (7.5 or 15 ng/ μL) of PCR product respectively and was made up to 10 μL with distilled water and was carried out at room temperature ($\sim 20^\circ\text{C}$) for 3 hours.

$$\text{Volume of insert for 1:1} = \frac{\left(\frac{([\text{Vector}] \times \text{Kb of Insert})}{\text{kb of Vector}} \right)}{[\text{Insert}]}$$

Equation 7-1: The volume calculation of 1 μL of the vector with the insert required for a 1:1 molar ratio. Where [Vector] is the concentration in ng/ μL and kb of vector and insert is the size of the DNA in kilobases.

7.4.4. Transformation

The Inoue method was followed for the preparation of competent cells.¹²¹ The transformation was carried out by adding 2 μ L ligation mixture into 50 μ L of competent NovaBlue (DE3) *E. coli* cells (Novagen) in a 1.5 mL sterile polypropylene tube (Eppendorf). The mixture was initially incubated at 4 °C for 10 min using an ice bath, followed by heat shock at 42 °C for 45 sec, and then placed on ice for a further 10 mins. Fresh Lennox Luria-Bertani media (800 μ L) was added to the mixture, followed by incubation at 37 °C with 180 rpm agitation for 1 h. Fresh LB medium (Lennox, Fischer Scientific, UK) contained 1 g of tryptone, 0.5 of yeast extract, 0.5 g of sodium chloride was dissolved in 100 mL of deionised water and autoclaved. Bacterial suspension aliquots (20 – 200 μ L) were plated onto LB agar plates with the relevant antibiotic such as (kanamycin (50 μ g/mL), chloramphenicol (35 μ g/mL), or ampicillin (100 μ g/mL)) as appropriate to facilitate selection and incubated at 37 °C for 16 h, to allow colony growth. The plates were then transferred to the fridge (4 °C) until used.

7.4.5. Cracking cells

Cracking was performed to confirm the PCR product had inserted into the vector. Up to eight colonies were picked from the LB agar plates using a sterile toothpick (Sigma Aldrich, UK). Each colony was inoculated into 5 mL of Luria-Bertani (LB) broth (Lennox) medium containing the appropriate antibiotic, by using conical centrifuge tubes (500/CS, 50 mL) and incubated at 37 °C, with shaking at 180 RPM for 16 hours. The next day, 20 μ L of each culture was mixed with 6 μ L cracking buffer containing: (50 % (w/v) sucrose, 10 % (w/v) SDS, 50 mM NaOH, and 25-

50 ng bromophenol blue in H₂O) in a 1.5 mL sterile polypropylene tube (Eppendorf). The samples were then loaded onto 1 % agarose gels alongside the corresponding undigested vector for comparison.

Plasmids of interest were then extracted using the miniprep extraction kit (Sigma-Aldrich, UK) after pelleting the cells by centrifugation at $2370 \times g$ (4600 RPM) for 10 min at 4 °C in a microcentrifuge, a fixed-angle rotor 24 x 1,5/2 mL 7500 3328, Biofuge Pico (Heraeus, Germany). Samples chosen for DNA extraction were subjected to the same restriction enzyme digest as used for cloning for 2 hours at room temperature (Section 7.4.1) to ensure that the PCR product had successfully ligated into the vector. Subsequent digested products were run on a 1 % agarose gels alongside a 100 bp, 1kb ladder and compared for size. Samples identified to contain an insert at the correct size were sent for sequencing at Source Bioscience (Nottingham).

7.4.6. DNA sequencing

DNA sequencing was performed on recombinant plasmids by the DNA Sequencing Facility at the University of Nottingham. 10 µL (90-110 µM) of plasmid DNA was sent for Sanger sequencing (Source Bioscience) to establish whether they contained the correct DNA sequence. Concentrations of plasmid DNA (ng/µL) were quantified using an ND-1000 Spectrophotometer NanoDrop, measuring the absorbance at 260 nm as described in Section 7.12.1. Confirmed target sequences were retransformed into BL21-CodonPlus (DE3)-RIL competent cells (Agilent Technologies) or LOBSTR BL21 (DE3)-RIL (Kerafast) cells for expression purposes.

7.4.7. Site-directed mutagenesis

FL-PRMT2 mutant was generated from native plasmids using Quick-change site-directed mutagenesis from Stratagene.¹³⁵ Primers encoding for an I401R mutation in PRMT2 (Table 7-6) were designed. Site-directed mutagenesis was performed to return to the original sequence of PRMT2. Mutation sites of the plasmids were confirmed by DNA sequencing as described previously.

Used for constructing	Forward primer 5' to 3'	Reverse primer 5' to 3'
FL-PRMT2 original sequence	ATGGCAACATCATATGACTGT	AAAGTCTTCCCCATCTGGAGA
FL-PRMT2 mutant sequence	ATGGCAACATCAGGTGACTGT	AAAGTCTTCCCCATCTGGAGA

Table 7-6: Primers used for the construction of an FL-PRMT2: I401R mutant using site-directed mutagenesis.

The PCR reaction contained 0.5 μ L (25-50 ng) of template DNA, 2 μ L (0.8 μ M) of forward primer, 2 μ L (0.8 μ M) of reverse primer, 0.2 μ L (200 μ M) of dNTPs (Thermo Fisher Scientific), 2 μ L (2 % v/v) of DMSO, 0.5 μ L (1 U) of PfuUltra II (Agilent Technologies) and 5 μ L (1X) of 10 $\times$ PfuUltra II buffer (Agilent Technologies) in a final volume of 50 μ L made up with sterile distilled water. The PCR conditions used are described in Table 7-7.

PCR process	Temperature	Time	cycles
Initial denaturation	95 °C	3 minutes	18
Denaturation	95 °C	1 minute	
Annealing	50 °C	1 minute	
Extension	68 °C	12 minutes	
Final extension	68 °C	30 minutes	

Table 7-7: PCR protocol used for FL-PRMT2 site direction mutagenesis.

After the PCR reaction was complete, 1 μ L (10 U) of *DpnI* (NEB) was added to the mixture and incubating for 3 hours at room temperature, to facilitate the digestion of the methylated template. Finally, the reaction mixture was then transformed into competent Nova Blue cells (Novagen). The transformation was performed as described in section 7.4.4. Plasmid DNA was subsequently extracted from the obtained colonies and extracted using a GenElute Plasmid Miniprep Kit (Sigma-Aldrich, UK). The purified sample was then sent for Sanger sequencing at Source Bioscience (Nottingham) to confirm the presence of the mutation, as previously described (Section 7.4.6).

7.5. BL21-CodonPlus (DE3)-RIL competent cells (Agilent), LOBSTR BL21 (DE3)-RIL and, Arctic Express competent cells

Various *E. coli* strains such as BL21-CodonPlus (DE3)-RIL competent cells (Agilent Technologies), Arctic Express competent cells (Agilent) and LOBSTR BL21 (DE3)-RIL (Kerafast) cells ⁸⁴ (see below), were used for protein production.

LOBSTR-BL21(DE3)-RIL or LOBSTR (low-background-strain) was derived from DE3 BL21cp, to increase the purity of the target protein.⁸⁴ LOBSTR-BL21 (DE3)-RIL is ideal for a low-copy protein target, and the cells maintained normal growth in the media like DE3 BL21-CodonPlus competent cells (Agilent). ArnA and SlyD were modified in this strain to reduce their protein affinity to the nickel chromatography column and in turn, improve the purity of the target protein (the wild type proteins contain several non-consecutive histidines and are the most common contaminant present in *E. coli* preparations).⁸⁴

Arctic Express competent cells were used to produce and purify PRMT6 with the aim of improving the yield of recombinant protein at low temperatures.^{106,105} This strain was chosen because *E. coli* Arctic Express (Stratagene) contains the chaperonins Cpn60/10 from a psychrophilic bacterium *Oleispira antarctica* and the presence of these chaperonins have been shown to facilitate proper protein folding.¹⁰⁵ As discussed in chapter 4, chaperonins were found to bind to the target protein, leading to co-purification, as evidenced by mass spectrometry (chapter 4). Chaperonin was separated from proteins by single step by treatment with a buffer consisting of (150 mM NaCl, 10 mM MgCl₂, 5 mM ATP and 20 mM HEPES-NaOH (pH 7.0)) at 4°C for 2 hours as discussed in more details in Section 7.8.3.

7.6. Gene expression and protein production

Once all sequences were confirmed and successfully cloned into the appropriate vector (see Section 7.3.6), they were transformed into BL21-CodonPlus (DE3)-RIL competent cells (Agilent), LOBSTR BL21 (DE3)-RIL cells or Arctic Express competent cells (Agilent) for the purpose of expressing the gene and for subsequent protein production. Transformation of each construct into these competent cells followed the same general protocol as used for transformation (see Section 7.4.4). To initiate expression, cells were taken from glycerol stocks (Section 7.3.5) or single colonies from agar plates and placed into 100 mL LB culture, which contained the appropriate antibiotic and incubated at 37 °C, with shaking at 180 rpm for 16 hours. Next day, 8 mL of a 100 mL culture was used to inoculate 800 mL of (LB) broth supplemented with 800 µL of the appropriate antibiotic (Table 7-8). LB medium (Lennox, Fischer Scientific, UK) contained 10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride was dissolved in 1-litre of

deionised water and autoclaved. The concentrations of antibiotic are shown in Table 7-8. Cells were grown at 37 °C, with shaking at 180 rpm, with an optical density at 600 nm on a spectrophotometer (OD₆₀₀) measured, until an OD₆₀₀ of 0.5–0.6 was reached. Upon reaching this optical density, 1.0 M isopropyl β-D-1-thiogalactopyranoside (IPTG, Thermo Scientific, Waltham, MA, USA) was added to a final concentration of 0.4 mM to the culture. The temperature was then lowered to the required temperature as summarised in (Table 7-8). After 16–24 h, cells were harvested by centrifugation at 2370 ×g (4600 RPM) for 30 min, at 4 °C in Rotanta 460 /460 R, rotor 4×1000 g (Hettich Centrifuges, UK). After pelleting the cells, the supernatant was removed, and the pellets were either used immediately for the next step as described below or stored at -20 °C until needed. Note, supernatants were not treated with benzonase, or subject to precipitation with either streptomycin sulfate or PEI because no remaining DNA was present as judged by the measured A280/A260 ratio).^{122,136}

chapter	Name of Construct	Expression Vector	Media	Antibiotics	Temperature after induction	Length of expression
3	PRMT1	pET26b	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	18-20 °C	18 hours
	MBP-CARM1	pMAL-c2xE	LB	Ampicillin (100 µg/mL) Chloramphenicol (35 µg/mL)	18-20 °C	18 hours
	CARM1-2GKG	pRSF-13	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	18-20 °C	18 hours
4	PRMT6	pPROEX HTb	LB	Ampicillin (100 µg/mL) Chloramphenicol (35 µg/mL)	15-20 °C	18 hours
	MBP-PRMT6	pMAL-c2XeE	LB	Ampicillin (100 µg/mL) Chloramphenicol (35 µg/mL)	15-20 °C	18 hours
	2GKG-PRMT6 1BKR-PRMT6	pRSF-13	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	15-20 °C	18 hours
	1BKR-PRMT6	pRSF-13	Auto induction	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	10-12 °C	72 hours
	H3	pET21a	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	18 °C	18 hours
5	FL-PRMT2	pColdI	LB	Ampicillin (100 µg/mL) Chloramphenicol (35 µg/mL)	10 °C	72 hours
	MBP-PRMT2	pMAL-c2x	LB	Ampicillin (100 µg/mL) Chloramphenicol (35 µg/mL)	10 °C	18 hours
	2GKG-PRMT2 1BKR-PRMT2	pRSF-13	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	15-20 °C	18 hours
	RSF1	pET21a	LB	Kanamycin (50 µg/mL) Chloramphenicol (35 µg/mL)	18 °C	18 hours

Table 7-8: Summary of different culture media and the lengths of expression for each construct.

7.7. Protein purification

7.7.1. General protein purification by FPLC

All chromatography was carried out using an ÄKTA FPLC systems or Akta PrimePlus (GE Healthcare Life Sciences, USA). The nickel affinity chromatography (prepacked 5 mL HisTrap HP) and Superdex 200 or 75 prep grade (HiLoad 16/600 Superdex 200 or 75 pg) columns were purchased from GE Healthcare. Purification was carried out at room temperature. Fractions of interest from the nickel column or the gel filtration were pooled and concentrated to 1-3 mL using a Vivaspin 20 centrifugal concentrator with a relevant molecular weight cut-off from GE Healthcare. The protein content of fractions from chromatography was analysed by SDS-PAGE (Section 7.9.1). All methods were performed on ice, and samples between each step were stored at 4 °C. All buffers for both the nickel column and the gel filtration were filtered using 47 mm, 0.2 µm pore sized Whatman filters (Sigma-Aldrich, UK) and degassed before use. The buffers routinely used for each construct are shown in Table 7-9.

7.7.2. Cell lysis

The resulting cell pellets were resuspended in 10 mL of buffer **A** containing: (300 mM NaCl, 10 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0) for each 1 g of cell pellet in 50 mL falcon tubes. Cell pellets in buffer **A** were distributed and subject to sonication (Hielscher UP200S Hielscher, Germany) with pulse-rest cycles of 20 seconds each, for a total of 4 minutes at 4 °C. Both sonication and resuspension were carried out on the ice. The resulting lysates were centrifuged in Heraeus Biofuge Stratos, 8 x 50 fixed angle rotor, R_{max} 10.7 mm, 34° (Thermo Scientific,

Heraeus Centrifuges, UK) at 14500-x-g (11,378 RPM) for 45 min, at 4 °C.; the supernatant was filtered using a 0.45 µm syringe filter (Sartorius) at 4 °C and loaded onto a 5 mL HisTrap HP column (GE Healthcare) as described below.

7.7.3. Nickel affinity chromatography

Nickel affinity chromatography (5 mL HisTrap HP column, GE Healthcare) was used to purify proteins bearing a 6×His tag using an ÄKTA Prime (GE Healthcare). Initially, 5 mL HisTrap HP column was pre-equilibrated with buffer **A**, using a peristaltic pump. The resulting supernatant was then loaded onto a pre-equilibrated 5 mL HisTrap HP column (GE Healthcare) and washed with 40-50 mL of buffer **A** (Table 7-9) removing any non-bound material. Proteins bound to the HisTrap HP column were subsequently eluted with an imidazole-HCl linear concentration gradient of between 20 mM to 500 mM in buffer **B** (Table 7-9) using ÄKTA Prime (GE Healthcare). Fractions of 2.5 mL were collected throughout this gradient elution. A small amount of protein from each fraction was analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) against a molecular weight ladder (Thermo Fisher Scientific). Fractions containing a protein with a molecular weight of interest were pooled and concentrated to 1-3 mL using Vivaspin 20 centrifugal concentrator with an appropriate molecular weight cut-off (Sigma-Aldrich, UK). The concentrated protein was then subject to the gel filtration, as described below.

7.7.4. Gel filtration

Preparative gel filtration was performed using HiLoad Superdex 16/60 columns (GE Healthcare). HiLoad Superdex 75 prep grade (pg) resin was used for the gel filtration of the histone and repressor splicing factor 1 (RSF1) peptides. While HiLoad Superdex 200 pg preppacked columns were used for all other proteins (Table 7-9). Columns were initially pre-equilibrated with the gel filtration buffer (Table 7-9). The fractions with target protein from the previous chromatographic step were centrifuged for 10 min at 12,000- $\times$ g (10,350 RPM) at 4 °C in a 6-place fixed-angle rotor (F-35-6-30) (centrifuged sigma, Eppendorf, Germany). Samples were then injected onto a pre-equilibrated the gel filtration HiLoad 16/16 column (Superdex™ 200 prep grad GE Healthcare) by using an injection loop (1 – 10 mL depending on sample volume). After the gradient elution was completed, a small amount of protein from each fraction was analysed by SDS-PAGE against a molecular weight ladder (gel filtration calibration kit GE healthcare). Fractions containing proteins were combined and concentrated using a Vivaspinn 10, 20 kDa MWCO centrifugal filter (Sigma-Aldrich, UK) and immediately used for purposes assays, or stored at -80 °C.

Chapter	Name of Construct	Purification Tag	Column used	Buffers Used
3	PRMT1 CARM1 CARM1-2GKG	C-term His	HiTrap chelating HP column pre-charged with the nickel (5 mL HisTrap HP column (GE Healthcare))	A: 300 mM NaCl, 5-10 % (v/v) glycerol, 20 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0 B: A with 500 mM imidazole-HCl
			HiLoad 16/600 Superdex 200 pg	300 mM NaCl, 2-3 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0
4	PRMT6-MBP PRMT6-2GKG PRMT6-1BKR	C-term His	HiTrap chelating HP column pre-charged with the nickel (5 mL HisTrap HP column (GE Healthcare))	A: 150-500 mM NaCl, 5-10 % (v/v) glycerol, 20 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0 B: A with 500 mM imidazole-HCl
5	PRMT2-MBP PRMT2-2GKG PRMT2-1BKR		HiLoad 16/600 Superdex 200 pg	150-500 mM NaCl, 2-3 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0
4 and 5	Histones H3	C-term His	HiTrap chelating HP column pre-charged with the nickel (5 mL HisTrap HP column (GE Healthcare))	A: 100 mM NaCl, 5-10 % (v/v) glycerol, 20 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0 B: A with 500 mM imidazole-HCl
5	Repressor splicing factor 1 (RSF1)	C-term His	HiLoad 16/600 Superdex S75 pg	100 mM NaCl, 2-3 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0

Table 7-9: The table shows chromatographic methods used for purification of a specific protein.

7.8. High-density shaking cultures by auto-induction media

7.8.1. Media preparation

PRMT6 was produced using autoinduction media with the aim of increasing yield and protein solubility. Protein production can be achieved by growing cells in minimal media, in the presence of a variable source, as described below. 1 L minimal auto-induction media was prepared in a 2 L non-baffled Erlenmeyer flask using a mixture of autoclaved reagents (Table 7-10) and a filter-sterilised sugar solution (Table 7-11). Initially, a combination of autoclaved reagents (solution 1) containing N-Z amine 10 g, Yeast extracts 5 g, and 2 mL of MgSO_4 (Table 7-10) was mixed with (solution 2) contained 25 mL of Na_2HPO_4 , 25 mL of KH_2PO_4 , 50 mL of NH_4Cl , 5 mL of NaSO_4 of 1.0 M stock solution (Table 7-10) in a 2 L non-baffled Erlenmeyer flask. The final addition was solution 3 (Table 7-11) which contained 20 mL glucose (1.0 % w/v), 20 mL lactose (8.0 % w/v) and 127 mL glycerol (100 % (v/v)) which was added to the mixture.

Autoclave	
Media preparation in 1 L (Solution 1)	Stock Salt solutions adjusted to pH 7.5 before adding to the mixture (Solution 2)
N-Z amine 10 g	1.0 M Na_2HPO_4
Yeast extract 5 g	1.0 M KH_2PO_4
2 mL of 1.0 M MgSO_4	1.0 M NH_4Cl
	1.0 M NaSO_4

Table 7-10: Individual components used to make minimal media and stock salt solution.

Filter sterilise sugar (solution 3)	
100 % (v/v) Glycerol	127 mL
1.0 % (w/v) Glucose	20 mL
8.0 % (w/v) Lactose	20 mL

Table 7-11: Individual components used to make minimal media. Glucose and lactose were filtered using a 0.2 μ m syringe filter.

Solutions 1 contained high concentrations of a variety of amino acids; as the presence of amino acids can increase the growth rate of cells and also influence the saturation density of glucose or glycerol.¹⁰⁶ Solution 2 was added to mediate metabolic acid production, maximise protein synthesis and soluble protein yield. Filter sterilised sugar solution 3 is a mixture of carbon sources which has been optimised for tightly regulating uninduced growth to high cell density.¹²³ Glucose in this setting can be considered an early energy source and prevents induction by lactose. However, glycerol is a late energy source and cells grow too much higher cellular densities before and after induction than that observed with lactose.¹²³

7.8.2. Gene expression and cell lysis

1BKR-PRMT6 (Chapter 4) was produced using *E. coli* Arctic Express cells by using a single colony or glycerol stock to inoculate 100 mL of autoclaved LB media (Lennox) at 37 °C for 16 hours with the appropriate antibiotic. 10 mL of the overnight culture was transferred into 1000 mL autoinduction media with the appropriate antibiotic (Table 7-8). The cultures were then incubated at 37 °C with continuous shaking at 250 rpm until an optical density (OD) between 0.6-0.8 was

reached. Upon reaching this optical density, the temperature was then lowered to 12 °C for 72 hours.¹⁰⁶ After 68 – 72 hours, cells were harvested by centrifugation at $2071 \times g$ (4,300 RPM) for 30 min at 4 °C in Rotanta 460/460 R, rotor 4×1000 g (Hettich Centrifuges, UK), used immediately followed instruction described in Section 7.7.2 or stored at -20 °C until needed.

7.8.3. A single step separation of co-purified chaperonins

Cell pellets containing the relevant PRMT and co-purified chaperonins from the previous expression step (Section 7.8.2) were resuspended in lysis buffer **A** containing (300 mM NaCl, 10 % (v/v) glycerol, 10 mM imidazole-HCl, 50 mM Tris-HCl and pH 8.0). The suspensions were sonicated and concentrated, as described in Section 7.7.2. The supernatant was then loaded onto a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with buffer **A**.

Separation of the co-purified chaperonin was achieved by initial washing with 100 mL of buffer **A**, then 150 mL of a buffer consisting of 150 mM NaCl, 10 mM MgCl₂, 5 mM ATP, 20 mM and HEPES-NaOH (pH 7.0) at 4 °C for 2 hours, using a peristaltic pump. After one hour, a second wash was performed as above. Finally, the resin was washed with 40 mL of buffer **A** to ensure sufficient removal of co-purified chaperonin. The bound proteins were subsequently eluted by an imidazole-HCl gradient of 20 mM to 500 mM, in (300 mM NaCl, 10 % (v/v) glycerol, 50 mM Tris-HCl and pH 8.0) over 25 column volumes on ÄKTA prime instrument. Combined fractions from the nickel affinity chromatography were pooled, and the gel filtration carried out using HiLoad Superdex 200 pg (GE Healthcare) as described in Section 7.7.4.

7.9. Protein analytical techniques

7.9.1. SDS-PAGE

SDS-PAGE analysis was used to estimate protein identity and purity. This technique involves the coating of proteins with negatively charged sodium dodecyl sulfate (SDS) and subsequent size separation through the application of an electrical current.^{124,134}

7.9.2. Gel preparation

SDS-PAGE gels with 0.75 mm thickness were prepared. Gels are comprised of two components, a stacking gel set on the top and resolving gel placed on the bottom and dimensions 10 × 10 cm were used. 15 % SDS-PAGE gels were mostly used (Table 7-12). For the resolving gel, 200 µL of isopropanol was added to the gel bed to ensure the gel set evenly for 20 min, and the isopropanol then discarded. For the top stacking gel, 15 well combs were added to the stacking solution prior to setting. Once set, the gel was assembled in the gel electrophoresis tank and filled with SDS Running Buffer (4 % (w/v) Glycine, 0.1 % (w/v) SDS, 125 mM Tris-HCl and pH 6.8). The well combs were carefully removed prior to protein sample loading.

Reagent	Resolving gel volume (mL)	Stacking gel volume (mL)
H ₂ O	3.7	5.4
30 % Acrylamide/Bis-Acrylamide (v/v)	10.6	2
1.5 M Tris-HCl pH 8.8	4	-
0.5 M Tris-HCl pH 6.8	-	2.5
10 % (w/v) SDS	0.16	0.1
10 % (w/v) APS	0.16	0.1
TEMED	0.016	0.01

Table 7-12: composition of 15 % SDS-PAGE gels

7.9.3. Sample preparation

The samples were prepared by mixing 5 μ L (10-25 μ g) of the protein of interest with 5 μ L of 2 $\times$ SDS loading dye sample buffer (160 mM Tris-HCl (pH 6.8), 5 % (w/v) SDS, 30 % (v/v) glycerol, 20 mM DTT, 0.03 % (w/v) bromophenol blue) in a sterile polypropylene tube. Before the samples were loaded onto the gels, the protein samples and the protein marker (NEB) were boiled for 5 minutes at 95 °C before loading into the wells. The first lane of each gel was loaded with a protein marker (broad range 2-212 kDa, NEB) for identifying the size of the protein. The gel was then run at 200 V using a BioRad Power Pack Basic for about 45 minutes until the band reached the bottom of the gel.

7.9.4. Gel staining

Gels were carefully transferred to a plastic container to stain with staining solution (25 % (v/v) isopropanol, 25 % (v/v) glacial acetic acid, 0.25 % (v/v) Coomassie Brilliant Blue R-250) for 30 min. The gel was then destained with an aqueous solution of 10 % (v/v) glacial acetic acid and 10 % (v/v) isopropanol in 1 L water for 30 min and then washed with water until protein bands became clearly visible. The gel was then visualised using a SynGene G: Box (Syngene, UK).

7.10. Western blot detecting His-tagged proteins

For western blot analysis, between 5 to 15 µg of each protein was used alongside positive and negative controls which were loaded onto two SDS-PAGE gels (one used for western blot and a second identical gel used for standard staining as shown in chapter 4 Section 4.4.1). The gel containing protein and all components of the transfer system were soaked with western blot transfer buffer (192 mM glycine, 25 mM Tris-HCl (pH 8.3), 20 % (v/v) methanol) prior to running the transfer. The gel was then transferred onto a nitrocellulose membrane (pre-soaked with MeOH) using a Mini Trans-Blot cell (Biorad), at 100 mA for 30 mins, at 4 °C. The nitrocellulose membrane was blocked for 16 h at 4 °C by incubating with blocking buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1 % (v/v) TWEEN-20, 5 % (w/v) skimmed milk powder).

Next day, the membrane was incubated with primary antibody (mouse IgG anti-tetra-His (Qiagen, product No: 34670) diluted 1/1000 in blocking buffer for 1 hour. After that, the blocking buffer was decanted and washed with TBST solution (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1 % (v/v) TWEEN-20) for 45 mins. This

process of washing was repeated three times using fresh TBST for 15 mins each to remove non-specific binding. The nitrocellulose membrane was incubated with a 1/1000 solution of the secondary antibody (Anti-Mouse IgG (Fab specific)– Peroxidase antibody (Sigma-Aldrich, product No: SAB4200791)) in blocking buffer for 2 hours at 4 °C. Again, the membrane was washed with TBST 20 mL for another 3 x 15-minutes.

A clean container was used to transfer the nitrocellulose membrane, followed by incubation for 1 min with ECL buffer (22 µL coumarin, 50 µL luminol and 3 µL hydrogen peroxide solution 30 % (w/w) in H₂O made up to 10 mL with 0.1 M Tris-HCl pH 8.5). Chemiluminescence was detected using an Image Quant LAS-4000 (GE Healthcare). Transfer of proteins was confirmed by Ponceau S red stain (Sigma-Aldrich, UK) to reveal the protein bands. All proteins were successfully detected on the membrane and compared to the second SDS-PAGE gel stained with Coomassie Blue to verify that all the proteins were successfully transferred to the membrane, as shown in the results section.

7.11. Crystallisation

Sitting drop crystallisation plates from Molecular Dimensions in conjunction with crystallisation screens were purchased from Molecular Dimensions or Hampton Research (Section 7.11.2 and 7.11.3). Crystallisation trials were performed by using the sitting drop vapour diffusion method in an MRC2 96-well plate. Numerous buffers and precipitants were used in crystallisation trials containing 75 or 80 µL in the reservoir well and 400 nL of protein in the sample well. Resulting crystals were cryoprotected in the drop condition with the addition of 20 % (v/v)

glycerol and cryo-cooling in liquid nitrogen prior to data collection. Several proteins with ligand were set up in crystallisation trials using conditions summarised in Table 7-13.^{125,126}

Chapter	Name of Construct with Ligand	Residue range	Concentration (mg/mL)	Crystallisation screens used
3	PRMT1+SAH CARM1+Compounds 55 and 58	22-361	3.44	PACT Screen JCSG plus
	CARM1+SAH CARM1+Compounds 46, 55, 58 , and 39	135-479	2.20 & 3.0	PACT Screen
	CARM1-2GKG	135-479	7.8 & 4.0	PACT Screen JCSG plus
4	PRMT6-1BKR	39-375	2.07 & 3.3	PACT Screen JCSG plus
5	PRMT2-2GKG PRMT2-1BKR	95-433	2.5 & 3.02	PACT Screen JCSG plus Morpheus

Table 7-13: Summary of proteins with ligands in crystallisation trials using various protein concentrations and crystallisation screens as specified. All crystallisation trials were conducted at 19 °C

7.11.1. Sample preparation and crystallisation

Gel filtration was initially performed to obtain pure protein and resulting target fractions concentrated to a final concentration of 2-5 mg/mL using a Vivaspin 3 kDa MWCO spin concentrator (GE Healthcare). Protein samples were centrifuged in a 6-place fixed-angle rotor (F-35-6-30) (Eppendorf, Germany) at 12,000 × g (10,350 RPM) for 10 min at 4 °C, in order to remove any debris or precipitated

protein. Proteins in complex with a ligand of interest were prepared with the 10-fold molar excess of ligand prior to being subject to crystallisation trials¹²⁶.

7.11.2. 96-well screens

MRC crystallisation plates (Molecular Dimensions) containing 96 well were set up using crystallisation screens from Hampton Research or Molecular Dimensions (Table 7-13). Each reservoir was loaded with 75 μ L of the precipitant solution using a multi-channel pipette. To form the sitting drops, 400 nL of each precipitant solution from each reservoir was transferred onto the corresponding shallow wells. An equal volume of protein solution was then added to the drop in each shallow well using a Mosquito Crystal liquid handler (TTP labtech). Trays were then sealed using HD Clear Duck packaging tape, and crystallisation plates left at 19 °C.¹²⁶

7.11.3. Optimisation screens using 24-well plates

Optimisation screens were carried out in attempts to grow higher quality crystals using 24 well-sitting drop VDX plates (Hampton Research). Each reservoir was loaded with 1 mL of precipitant solution. 1 μ L of precipitant solution from each reservoir was transferred onto its corresponding shallow well. An equal volume of protein in complexes was mixed in each drop shallow well. The plate was then sealed and left at 19 °C to grow the crystals.

7.11.4. Crystal fishing and freezing

Crystals were fished using MicroLoops (MiTeGen) and transferred carefully into 5 μL of cryoprotectant buffer comprised of the specific crystallisation buffer with 25 % (v/v) of glycerol or ethylene glycol. Crystals were then immediately snap-frozen and stored in liquid nitrogen ready for shipping.^{127,128}

7.12. Biophysical techniques

7.12.1. UV spectroscopy

Protein and DNA concentrations were determined from UV absorbance using the Beer-Lambert equation.

$$A = \epsilon cl$$

The Beer-Lambert equation, where, A is absorbance in AU; ϵ is the molar extinction coefficient in $\text{M}^{-1}\text{cm}^{-1}$, c is a concentration in M , and l is the path length in cm .

Molar extinction coefficients (ϵ) were generally taken from the ExPASy ProtParam online tool.¹³³ Bisubstrate inhibitors and ligand concentration were determined using a published extinction coefficient for S-adenosyl homocysteine determined at 260 nm ($\epsilon^{260} = 15400 \text{ M}^{-1}\text{cm}^{-1}$).¹²⁹

A ND-1000 NanoDrop spectrophotometer (Thermo Scientific) was used to measure concentrations of ligand, plasmid DNA, and protein samples, respectively. UV absorbance was measured at the wavelength specified below for 2 μL of ligand, plasmid DNA, or protein obtained directly from the gel filtration buffer compared to the same volume of 'blank' buffer. The absorbance of protein samples was

measured at 280 nm, While ligand and plasmid DNA concentrations were measured at 260 nm..

7.12.2. Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) was performed to assess binding between protein and ligand using a MicroCal PEAQ-ITC¹¹⁰(Malvern Panalytical). Fresh protein was taken from the gel filtration and dialysed into the buffer (150-300 mM NaCl, 1 % (v/v) glycerol, 20 mM Tris-HCl and pH 8.0) as suitable for ITC titrations. The sample protein was then concentrated to 20-35 μ M as determined by UV absorbance and loaded into the sample cell. Various concentrations of protein were tested to achieve sufficient curve sigmoidicity. The ligands were diluted into the same buffer as used for the protein and concentrated to 200-350 μ M and approximately 75 μ L loaded into the microsyringe. Ligands were titrated into the sample cell in 17 x 2.4 μ L aliquots at intervals 180 seconds apart and a stirring speed of 1000 rpm, at 25°C. Control experiments determined the heat of dilution of the ligand by injecting the same ligand solutions to the same buffer in the absence of protein. The control of ligand to buffer was subtracted from binding experiments, using a one-site binding isotherm, with fitted offset varied. The concentrations of protein and ligand used are documented in Table 7-14 (below), and the corresponding curves are given in the result and discussion chapters. The resulting data was analysed by curve fitting using the competitive binding model in the MicroCal PEAQ-ITC Analysis Software.^{110,130}

Chapter	Protein	Inhibitors	[protein] (cell)/ μM	[Ligands] (syringe)/ μM
3	PRMT1	SAH 55 58 46	30	300
3	CARM1	SAH 55 58 46	20-30	200-300
3	2GKG- CARM1	SAH	20	200
4	1BKR- PRMT6	SAH	29	290
5	1BKR- PRMT2	SAH	20, 30, and 40	200,300, and 400

Table 7-14: Concentrations of Protein and ligands used for isothermal titration calorimetry.

8. Chapter 8 References

- (1) Shen, Y.; Szewczyk, M. M.; Eram, M. S.; Smil, D.; Kaniskan, H. Ü.; Ferreira De Freitas, R.; Senisterra, G.; Li, F.; Schapira, M.; Brown, P. J.; *et al.*, *J. Med. Chem.* **2016**, 59, 9124–9139.
- (2) Muhsen, U. A. Toward the Development of Novel and Selective Bisubstrate Inhibitors of Protein Arginine Methyltransferases. PhD Thesis, University of Nottingham, **2014**.
- (3) Hong, W. Design and Synthesis of Protein Arginine Methyltransferase Inhibitors. PhD Thesis University of Nottingham, **2010**.
- (4) Copeland, R. A.; Solomon, M. E.; Richon, V. M. *Nat. Rev. Drug Discov.* **2009**, 8, 724–732.
- (5) Gayatri, S.; Bedford, M. T. *Biochim. Biophys. Acta-Gene Regul. Mech.* **2014**, 1839, 702–710.
- (6) Schluckebier, G.; O’Gara, M.; Saenger, W.; Cheng, X. *J. Mol. Biol.* **1995**, 247, 16–20.
- (7) Wolf, S. S. *Cell. Mol. Life Sci.* **2009**, 66, 2109–2121.
- (8) Cantoni, G. L. *J. Am. Chem. Soc.* **1952**, 74, 2942–2943.
- (9) Schubert, H. L.; Blumenthal, R. M.; Cheng, X. *Trends Biochem. Sci.* **2003**, 28, 329–335.
- (10) Fontecave, M.; Atta, M.; Mulliez, E. *Trends Biochem. Sci.* **2004**, 29, 243–249.
- (11) Tal, J.; Zeng, G. C.; Doctor, B. P.; Gordon, K. *FASEB J.* **1996**, 10, 471–480.
- (12) Stolowitz, M. L.; Minch, M. J. *J. Am. Chem. Soc.* **1981**, 103, 6015–6019.
- (13) Matos, J. R.; Wong, C. H. *Bioorg. Chem.* **1987**, 15, 71–80.
- (14) Hoffman, J. L. *Biochemistry.* **1986**, 25, 4444–4449.

- (15) Herdewijn, P. *Modified Nucleosides: in Biochemistry, Biotechnology and Medicine*, Wiley-VCH: Weinheim, **2008**.
- (16) Hobley, G.; McKelvie, J. C.; Harmer, J. E.; Howe, J.; Oyston, P. C. F.; Roach, P. L. *Bioorg Med. Chem. Lett.* **2012**, 22, 3079–3082.
- (17) Di Lorenzo, A.; Bedford, M. T. *FEBS Lett.* **2011**, 585, 2024–2031.
- (18) Rust, H. L.; Zurita-Lopez, C. I.; Clarke, S.; Thompson, P. R. *Biochemistry.* **2011**, 50, 3332–3345.
- (19) Tang, J.; Frankel, A.; Cook, R. J.; Kim, S.; Paik, W. K.; Williams, K. R.; Clarke, S.; Herschman, H. R. *J. Biol. Chem.* **2000**, 275, 7723–7730.
- (20) Fuhrmann, J.; Clancy, K. W.; Thompson, P. R. *Chem. Rev.* **2015**, 115, 5413–5461.
- (21) Haas, D. J.; Harris, D. R.; Mills, H. H. *Acta Crystallogr.* **1965**, 19, 676–679.
- (22) Hughes, R. M.; Waters, M. L. *J. Am. Chem. Soc.* **2006**, 128, 12735–12742.
- (23) Pahlich, S.; Zakaryan, R. P.; Gehring, H. *Biochim. Biophys. Acta - Proteins Proteomics.* **2006**, 1764, 1890–1903.
- (24) Zhang, X.; Cheng, X. *Structure.* **2003**, 11, 509–520.
- (25) Boriack-Sjodin, P. A.; Jin, L.; Jacques, S. L.; Drew, A.; Sneeringer, C.; Scott, M. P.; Moyer, M. P.; Ribich, S.; Moradei, O.; Copeland, R. A. *ACS Chem. Biol.* **2016**, 11, 763–771.
- (26) Osborne, T. C.; Obianyo, O.; Zhang, X.; Cheng, X.; Thompson, P. R. *Biochemistry.* **2007**, 46, 13370–13381.
- (27) Mitchell, L. H.; Drew, A. E.; Ribich, S. A.; Rioux, N.; Swinger, K. K.; Jacques, S. L.; Lingaraj, T.; Boriack-Sjodin, P. A.; Waters, N. J.; Wigle, T. J.; *et al.*, *ACS Med. Chem. Lett.* **2015**, 6, 655–659.
- (28) Bedford, M. T.; Richard, S. *Mol. Cell.* **2005**, 18, 263–272.
- (29) Bedford, M. T. *J. Cell Sci.* **2007**, 120, 4243–4246.

- (30) Qi, C.; Chang, J.; Zhu, Y.; Yeldandi, A. V.; Rao, S. M.; Zhu, Y. J. *J. Biol. Chem.* **2002**, 277, 28624–28630.
- (31) Cura, V.; Marechal, N.; Troffer-Charlier, N.; Strub, J. M.; van Haren, M. J.; Martin, N. I.; Cianferani, S.; Bonnefond, L.; Cavarelli, J. *FEBS J.* **2017**, 284, 77–96.
- (32) Bedford, M. T.; Clarke, S. G. *Mol. Cell.* **2009**, 33, 1–13.
- (33) McBride, A. E.; Silver, P. A. *Cell.* **2001**, 106, 5–8.
- (34) Bonnefond, L.; Stojko, J.; Mailliot, J.; Troffer-Charlier, N.; Cura, V.; Wurtz, J. M.; Cianf erani, S.; Cavarelli, J. *J. Struct. Biol.* **2015**, 191, 175–183.
- (35) Herrmann, F.; Pably, P.; Eckerich, C.; Bedford, M. T.; Fackelmayer, F. O. *J. Cell Sci.* **2009**, 122, 667–677.
- (36) Poulard, C.; Corbo, L.; Le Romancer, M. *Oncotarget.* **2016**, 7(41), 67532–67550.
- (37) Elakoum, R.; Gauchotte, G.; Oussalah, A.; Wissler, M. P.; Cl ement-Duch ene, C.; Vignaud, J. M.; Gu ant, J. L.; Namour, F. *Biochimie.* **2014**, 97, 210–218.
- (38) Yoshimatsu, M.; Toyokawa, G.; Hayami, S.; Unoki, M.; Tsunoda, T.; Field, H. I.; Kelly, J. D.; Neal, D. E.; Maehara, Y.; Ponder, B. A. J.; *et al.*, *Int. J. Cancer.* **2011**, 128, 562–573.
- (39) Xu, W.; Cho, H.; Kadam, S.; Banayo, E. M.; Anderson, S.; Yates, J. R.; Emerson, B. M.; Evans, R. M. *Genes Dev.* **2004**, 18, 144–156.
- (40) Drew, A. E.; Moradei, O.; Jacques, S. L.; Rioux, N.; Boriack-Sjodin, A. P.; Allain, C.; Scott, M. P.; Jin, L.; Raimondi, A.; Handler, J. L.; *et al.*, *Sci. Rep.* **2017**, 7, 1–13.
- (41) Sack, J. S.; Thieffine, S.; Bandiera, T.; Fasolini, M.; Duke, G. J.; Jayaraman, L.; Kish, K. F.; Klei, H. E.; Purandare, A. V.; Rosettani, P.; *et al.*, *Biochem.*

- J.* **2011**, 436, 331–339.
- (42) Troffer-Charlier, N.; Cura, V.; Hassenboehler, P.; Moras, D.; Cavarelli, J. *EMBO J.* **2007**, 26, 4391–4401.
- (43) Yoshimoto, T.; Boehm, M.; Olive, M.; Crook, M. F.; San, H.; Langenickel, T.; Nabel, E. G. *Exp. Cell Res.* **2006**, 312, 2040–2053.
- (44) Frankel, A.; Yadav, N.; Lee, J.; Branscombe, T. L.; Clarke, S.; Bedford, M. T. *J. Biol. Chem.* **2002**, 277, 3537–3543.
- (45) Halby, L.; Marechal, N.; Pechalrieu, D.; Cura, V.; Franchini, D. M.; Faux, C.; Alby, F.; Troffer-Charlier, N.; Kudithipudi, S.; Jeltsch, A.; *et al.*, *Philos. Trans. R. Soc. B: Biol. Sci.* **2018**, 373 (1748), 20170072.
- (46) Wang, C.; Zhu, Y.; Chen, J.; Li, X.; Peng, J.; Chen, J.; Zou, Y.; Zhang, Z.; Jin, H.; Yang, P.; *et al.*, *PLoS One.* **2014**, 9, 1–12.
- (47) Struck, A. W.; Thompson, M. L.; Wong, L. S.; Micklefield, J. *ChemBioChem.* **2012**, 13, 2642–2655.
- (48) Lee, D. Y.; Teyssier, C.; Strahl, B. D.; Stallcup, M. R. *Endocr. Rev.* **2005**, 26, 147–170.
- (49) Xiao, B.; Jing, C.; Wilson, J. R.; Walker, P. A.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I. A.; Blackburn, G. M.; Gamblin, S. J. *Nature.* **2003**, 421, 652–656.
- (50) Dillon, S. C.; Zhang, X.; Trievel, R. C.; Cheng, X. *Genome. Biol.* **2005**, 6, 227.
- (51) Luo, M. *ACS Chem. Biol.* **2012**, 7, 443–463.
- (52) Shaaban, S. A.; Bedford, M. T. *Chem. Biol.* **2007**, 14, 242–244.
- (53) Ragno, R.; Simeoni, S.; Castellano, S.; Vicidomini, C.; Mai, A.; Caroli, A.; Tramontano, A.; Bonaccini, C.; Trojer, P.; Bauer, I.; *et al.*, *J. Med. Chem.* **2007**, 50, 1241–1253.

- (54) Reeve, A. M.; Townsend, C. A. *Tetrahedron*. **1998**, *54*, 15959–15974.
- (55) Cheng, D.; Yadav, N.; King, R. W.; Swanson, M. S.; Weinstein, E. J.; Bedford, M. T. *J. Biol. Chem.* **2004**, *279*, 23892–23899.
- (56) Kaniskan, H. Ü.; Konze, K. D.; Jin, J. *J. Med. Chem.* **2015**, *58*, 1596–1629.
- (57) Liu, F.; Li, F.; Ma, A.; Dobrovetsky, E.; Dong, A.; Gao, C.; Korboukh, I.; Liu, J.; Smil, D.; Brown, P. J.; *et al.*, *J. Med. Chem.* **2013**, *56*, 2110–2124.
- (58) Eram, M. S.; Shen, Y.; Szewczyk, M. M.; Wu, H.; Senisterra, G.; Li, F.; Butler, K. V.; Kaniskan, H. Ü.; Speed, B. A.; Dela Seña, C.; *et al.*, *ACS Chem. Biol.* **2016**, *11*, 772–781.
- (59) Dowden, J.; Hong, W.; Parry, R. V.; Pike, R. A.; Ward, S. G. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2103–2105.
- (60) Osborne, T.; Roska, R. L. W.; Rajsiki, S. R.; Thompson, P. R. *J. Am. Chem. Soc.* **2008**, *130*, 4574–4575.
- (61) Dowden, J.; Pike, R. A.; Parry, R. V.; Hong, W.; Muhsen, U. A.; Ward, S. G. *Org. Biomol. Chem.* **2011**, *9*, 7814–7821.
- (62) Van Haren, M.; Van Ufford, L. Q.; Moret, E. E.; Martin, N. I. *Org. Biomol. Chem.* **2015**, *13*, 549–560.
- (63) Brooks, B. R.; III, C. L. B.; A. D. Mackerell, J.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; *et al.*, *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (64) Wang, R.; Islam, K.; Liu, Y.; Zheng, W.; Tang, H.; Lailier, N.; Blum, G.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2013**, *153*, 1048–1056.
- (65) Lin, Q.; Jiang, F.; Schultz, P. G.; Gray, N. S. *J. Am. Chem. Soc.* **2001**, *123*, 11608–11613.
- (66) Ramalingam, K.; Woodard, R. W. *J. Org. Chem.* **1988**, *53*, 1900–1903.
- (67) Halligan, K.; Nair, V. *Arkivoc.* **2006**, *2006*, 101–106.

- (68) Enschedé, C.; Hesse, M. *Helv. Chim. Acta.* **2002**, *85*, 1659–1673.
- (69) Huang, S. L.; Swern, D. J. *Org. Chem.* **1978**, *43*, 4537–4538.
- (70) Heckmann, D.; Meyer, A.; Laufer, B.; Zahn, G.; Stragies, R.; Kessler, H. *ChemBioChem.* **2008**, *9*, 1397–1407.
- (71) Gilchrist, T. L.; Moody, C. J. *Chem. Rev.* **1977**, *77*, 409–435.
- (72) PCT, I. *Appl.* **2009**, No. WO 2009111679A2.
- (73) Bergeron, R. J.; Mcmanis, J. S. *J. Org. Chem.* **1987**, *52*, 1700–1703.
- (74) Bothwell, I. R.; Islam, K.; Chen, Y.; Zheng, W.; Blum, G.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2012**, *134*, 14905–14912.
- (75) Guo, H.; Wang, R.; Zheng, W.; Chen, Y.; Blum, G.; Deng, H.; Luo, M. *ACS Chem. Biol.* **2014**, *9*, 476–484.
- (76) Wang, R.; Zheng, W.; Luo, M. *Anal. Biochem.* **2014**, *450*, 11–19.
- (77) Chong, Y.; Young, J.; Kim, J.; Lee, Y.; Park, K. S.; Cho, J. H.; Kwon, H. J.; Suh, J. W.; Lim, Y. *J. Microbiol. Biotechnol.* **2006**, *16*, 1154–1157.
- (78) Pignot, M.; Pljevaljcic, G.; Weinhold, E. *Eur. J. Org. Chem.* **2000**, *2000*, 549–555.
- (79) Scott, D. E.; Dawes, G. J.; Ando, M.; Abell, C.; Ciulli, A. *ChemBioChem* **2009**, *10*, 2772–2779.
- (80) Boyle, P. H.; Davis, A. P.; Dempsey, K. J.; Hosken, G. D. *Tetrahedron: Asymmetry.* **1995**, *6*, 2819–2828.
- (81) Meffre, P. *Amino Acids.* **1999**, *16*, 251–272.
- (82) Barton, D. H. R.; Géro, S. D.; Lawrence, F.; Robert-Gero, M.; Quiclet-Sire, B.; Samadi, M. *J. Med. Chem.* **1992**, *35*, 63–67.
- (83) Gunnell, E. Investigating Bisubstrate Inhibitors of PRMT1 and CARM1. PhD Thesis, University of Nottingham, **2018**.
- (84) Andersen, K. R.; Leksa, N. C.; Schwartz, T. U. *Proteins Struct. Funct.*

- Bioinforma.* **2013**, *81*, 1857–1861.
- (85) Strahl, B. D.; Briggs, S. D.; Brame, C. J.; Caldwell, J. A.; Koh, S. S.; Ma, H.; Cook, R. G.; Shabanowitz, J.; Hunt, D. F.; Stallcup, M. R.; *et al.*, *Curr. Biol.* **2001**, *11*, 996–1000.
- (86) Yue, W. W.; Hassler, M.; Roe, S. M.; Thompson-Vale, V.; Pearl, L. H. *EMBO J.* **2007**, *26*, 4402–4412.
- (87) Kuszewski, J.; Gronenborn, A. M.; Clore, G. M. *J. Am. Chem. Soc.* **1999**, *121*, 2337–2338.
- (88) Damian, L., Isothermal titration calorimetry for studying protein-ligand interactions. *Methods Mol. Biol.* **2013**, *1008*, 103–18.
- (89) Newman, J.; Egan, D.; Walter, T. S.; Meged, R.; Berry, I.; Jelloul, M. Ben; Sussman, J. L.; Stuart, D. I.; Perrakis, A. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, *61*, 1426–1431.
- (90) Kapust, R. B.; Waugh, D. S. *Protein Sci.* **1999**, *8*, 1668–1674.
- (91) Moon, A. F.; Mueller, G. A.; Zhong, X.; Pedersen, L. C. *Protein Sci.* **2010**, *19*, 901–913.
- (92) Reuten, R.; Nikodemus, D.; Oliveira, M. B.; Patel, T. R.; Brachvogel, B.; Breloy, I.; Stetefeld, J.; Koch, M. *PLoS One.* **2016**, *11*, 1–15.
- (93) Fraser, J. S.; Merlie, J. P.; Echols, N.; Weisfield, S. R.; Mignot, T.; Wemmer, D. E.; Zusman, D. R.; Alber, T. *Mol. Microbiol.* **2007**, *65*, 319–332.
- (94) Caulton, S. Characterisation of the Deubiquitinating Enzyme USP20. PhD Thesis, University of Nottingham, **2017**.
- (95) Qian, K.; Hu, H.; Xu, H.; Zheng, Y. G. *Signal Transduct. Target. Ther.* **2018**, *3*, (1), 6.
- (96) Van Haren, M. J.; Marechal, N.; Troffer-Charlier, N.; Cianciulli, A.; Sbardella, G.; Cavarelli, J.; Martin, N. I. *Proc. Natl. Acad. Sci.* **2017**, *114*, 3625–3630.

- (97) Dessau, M. A.; Modis, Y. J. *Vis. Exp.* **2011**, 9, 1–6.
- (98) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66, 486–501.
- (99) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; *et al.*, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66, 213–221.
- (100) Yang, Z. R.; Thomson, R.; McNeil, P.; Esnouf, R. M. *Bioinformatics.* **2005**, 21, 3369–3376.
- (101) Hewitt, S. N.; Choi, R.; Kelley, A.; Crowther, G. J.; Napuli, A. J.; Van Voorhis, W. C. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2011**, 67, 1006–1009.
- (102) Liu, Y.; Manna, A.; Li, R.; Martin, W. E.; Murphy, R. C.; Cheung, A. L.; Zhang, G. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 6877–6882.
- (103) Bañuelos, S.; Saraste, M.; Carugo, K. D. *Structure.* **1998**, 6, 1419–1431.
- (104) Thomas, D.; Lakowski, T. M.; Pak, M. L.; Kim, J. J.; Frankel, A. *Protein Sci.* **2010**, 19, 2141–2151.
- (105) Joseph, R. E.; Andreotti, A. H. *Protein Expr. Purif.* **2008**, 60, 194–197.
- (106) Studier, F. W. *Protein Expr. Purif.* **2005**, 41, 207–234.
- (107) Tuli, L.; W. Resson, H. J. *Proteomics Bioinform.* **2009**, 02, 416–438.
- (108) Lakowski, T. M.; Frankel, A. *Biochem. J.* **2009**, 421, 253–261.
- (109) Wallington, E. J.; Lund, P. A. *Microbiology.* **1994**, 140, 113–122.
- (110) Okhrimenko, O.; Jelesarov, I. J. *Mol. Recognit.* **2008**, 21, 1–19.
- (111) Iberg, A. N.; Espejo, A.; Cheng, D.; Kim, D.; Michaud-Levesque, J.; Richard, S.; Bedford, M. T. *J. Biol. Chem.* **2008**, 283, 3006–3010.
- (112) Heras, B.; Martin, J. L. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, 61, 1173–1180.

- (113) Bergfors, T. J. *Struct. Biol.* **2003**, 142, 66–76.
- (114) Krauss, I. R.; Sica, F.; Mattia, C. A.; Merlino, A. *Int. J. Mol. Sci.* **2012**, 13, 3782–3800.
- (115) Qing, G.; Ma, L. C.; Khorchid, A.; Swapna, G. V. T.; Mal, T. K.; Takayama, M. M.; Xia, B.; Phadtare, S.; Ke, H.; Acton, T.; *et al.*, *Nat. Biotechnol.* **2004**, 22, 877–882.
- (116) Hoffmann, A.; Bukau, B.; Kramer, G. *Biochim. Biophys. Acta - Mol. Cell Res.* **2010**, 1803, 650–661.
- (117) Costa, S.; Almeida, A.; Castro, A.; Domingues, L. *Front. Microbiol.* **2014**, 5, 1–20.
- (118) Li, Y. *Biotechnol. Appl. Biochem.* **2010**, 55, 73–83.
- (119) Blythe, S. A.; Cha, S. W.; Tadjuidje, E.; Heasman, J.; Klein, P. S. *Dev. Cell.* **2010**, 19, 220–231.
- (120) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; *et al.*, *Nucleic Acids Res.* **2014**, 42, 252–258.
- (121) Inoue, H.; Nojima, H.; Okayama, H. *Gene.* **1990**, 96 (1), 23–28..
- (122) Stanley, B. P.; Darlene, F. G.; David, C. B.; Stephen, B. K.; Leroy, E. H. *Cell.* **1984**, 38, 127–134
- (123) Grabski, A.; Spangler, N.; Ginos, M.; Drott, D.; Biosciences, E. M. D. *Nat. Methods.* **2005**, 2, 233–235.
- (124) George, A. J. T.; Urch, C. E.; Al-Tubuly, A. A. *Diagnostic Ther. Antibodies.* **2003**, 40, 391–405.
- (125) Chayen, N. E. *Curr. Opin. Struct. Biol.* **2004**, 14, 577–583.
- (126) Chung, C. W. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, 63, 62–71.
- (127) Garman, E. F.; Owen, R. L. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**,

- 62, 32–47.
- (128) Sliz, P.; Harrison, S. C.; Rosenbaum, G. *Structure*. **2003**, *11*, 13–19.
- (129) Van Laer, B.; Roovers, M.; Wauters, L.; Kasprzak, J. M.; Dyzma, M.; Deyaert, E.; Singh, R. K.; Feller, A.; Bujnicki, J. M.; Droogmans, L.; *et al.*, *Nucleic Acids Res.* **2016**, *44*, 940–953.
- (130) Du, X.; Li, Y.; Xia, Y. L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. *Int. J. Mol. Sci.* **2016**, *17*, 144.
- (131) Ealick, S. E.; Walter, R. L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 725–736.
- (132) Kirshenbaum, N.; Michaelevski, I.; Sharon, M. *J. Vis. Exp.* **2010**, *40*, 3–8.
- (133) ExPASy-ProtParam tool <https://web.expasy.org/protparam/> (accessed May 19, **2019**).
- (134) Laemmli, U. *Nature*. **1970**, *227*, 680–685.
- (135) http://kirschner.med.harvard.edu/files/protocols/Stratagene_quickchange.pdf (accessed Jun 8, **2019**).
- (136) Glasel, J. "Validity of nucleic acid purities monitored by 260/280 absorbance ratios". *BioTechniques*. **1995**, *18* (1): 62–63.
- (137) PyMOL Molecular Graphics System. <http://sourceforge.net/projects/pymol/> (accessed Sep 15, **2019**).

9. Appendix

Experimental methods performed by Dr Emma Gunnell

9.1. *In Vitro* radiometric methylation assay

Enzymatic experiments were carried out utilising the *in vitro* radiometric methylation assay. Each reaction was composed of recombinant PRMT (1 μ L) and the appropriate H4 or H3 histone peptide substrate (1 μ L) in the presence of [^{3} H]-SAM 0.5 μ L, (Perkin Elmer, NET155H250UC at a final concentration of 0.23 μ M) and PBS was used to make the total reaction volume up to 15 μ L with PBS. Recombinant PRMT and histone peptide in PBS were gently mixed in an Eppendorf tube and reactions were initiated by the addition of [^{3} H]-SAM at 30 °C for 1 hour. After one hour, the mixture reaction was spotted onto p81 paper. Once dried, the p81 paper was washed with a 10 % (w/v) aqueous solution of TCA (Trichloroacetic acid) for 45 min, and then was replaced with a fresh aqueous solution of TCA after every 15 min. After the TCA washes, the paper was washed with an ethanol solution and then left to dry. The dried paper was then transferred to a scintillation vial, and 2mL of scintillation fluid was added. A scintillation counter was used to determine the amount of [^{3} H]-SAM bound to the paper.⁴ The data obtained from the *in vitro* methylation assay was then plotted using the GraphPad Prism software.

Chapter	Construct	Concentration (μ M)	Substrate	Concentration (μ M)
3	PRMT1	160 nM	GB-H4 (1-21) peptide	40 μ M
3	CARM1	230 nM	GB-H3 (3-31) peptide	120 μ M
4	1BKR-PRMT6	7.0 μ M	GB-H3 (1-22) peptide	76 μ M
5	1BKR-PRMT2	27	GB-H3 (1-22) peptide GB-H4 (1-21) peptide GB-RSF1(1-22) peptide	39

Table 1: Enzyme and substrate concentrations used for initial activity tests of purified recombinant constructs.

9.2. X-Ray diffraction, data collection, structure determination and refinement

X-ray diffraction was carried out at the Diamond light source. Data from the protein crystals were collected at beamline I24 at various diffraction resolutions (described in Result and Discussion chapters). Data was processed utilising X-ray Detector Software (XDS) and scaled, merged utilising Aimless.¹³¹ The structure of proteins was solved by molecular replacement against the wild-type crystal structure of CARM1 (PDB 2Y1X).⁴⁰ Refinement was performed in PHENIX.⁹⁹ Refine by utilising Translation/Libration/Screw (TLS) groups. Non-crystallographic symmetry (NCS) restraints and the model were rebuilt in Coot. The solved structure files were then deposited at the protein data bank (PDB).